Nonlinear Dyn
https://doi.org/10.1007/s11071-023-08247-7

ORIGINAL PAPER

Sorting-free Hill-based stability analysis of periodic solutions through Koopman analysis

Fabia Bayer · Remco I. Leine

Received: 19 October 2022 / Accepted: 20 December 2022

Abstract In this paper, we aim to study nonlinear time-periodic systems using the Koopman operator, which provides a way to approximate the dynamics of a nonlinear system by a linear time-invariant system of higher order. We propose for the considered system class a specific choice of Koopman basis functions combining the Taylor and Fourier bases. This basis allows to recover all equations necessary to perform the harmonic balance method as well as the Hill analysis directly from the linear lifted dynamics. The key idea of this paper is using this lifted dynamics to formulate a new method to obtain stability information from the Hill matrix. The error-prone and computationally intense task known by *sorting*, which means identifying the best subset of approximate Floquet exponents from all available candidates, is circumvented in the proposed method. The Mathieu equation and an n-DOF generalization are used to exemplify these findings.

Keywords Time-periodic systems · Harmonic balance method · Koopman lift · Floquet multipliers · Monodromy matrix

F. Bayer (✉) · R. I. Leine
Institute for Nonlinear Mechanics, University of Stuttgart,
Pfaffenwaldring 9, 70569 Stuttgart, Germany
e-mail: bayer@inm.uni-stuttgart.de

R. I. Leine
e-mail: leine@inm.uni-stuttgart.de

1 Introduction

The objective of this paper is to introduce a novel stability method based on the Hill matrix, which differs from the state-of-the-art methods in that a matrix projection is applied before computing an eigenvalue problem. The structure of this projection is obtained by considering the Hill matrix to be a result from the Koopman lift for well-chosen basis functions.

The Koopman framework [1,2] has gained immense popularity in recent years as a versatile tool for various engineering applications, such as system identification [3], model order reduction [4] and feedback control [5]. This is due to an auspicious promise: global representation of a nonlinear system by a linear operator. To this end, in the Koopman framework, the dynamical system is defined through the propagation of functions on the state space, also called observables, over time. Bernard Koopman first described a unitary linear operator which evolves a class of measurable functions along the flow of a conservative system [6], which would later be named the Koopman operator. After some relatively quiet years, the Koopman operator experienced a revival around the turn of the century, when it was shown that its spectral characteristics contain global properties for the underlying dynamical system [7]. This sparked generalizations to non-conservative systems [8,9] and at the same time, numerical and data-driven methods like the Arnoldi method [10] and extended dynamic mode decomposition [2] emerged.

Published online: 06 February 2023

Classically, the Koopman framework is applied to time-autonomous systems $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and the approximate linear dynamics obtained by the Koopman lift then takes the form $\dot{\mathbf{z}} = \mathbf{A}\mathbf{z}$. The incorporation of a time-dependent input $\mathbf{v}(t)$ into the dynamics, i.e., $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \mathbf{v}(t))$ or simply $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$, generally poses problems in the Koopman framework as the system can only be approximated by a linear time-invariant (LTI) system $\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\mathbf{u}(t)$ if products of state and input are neglected. In this paper we focus on non-autonomous systems for which the input is time-periodic, i.e., $\mathbf{v}(t) = \mathbf{v}(t + T)$. In particular, we propose a specific choice of observable functions which contains observables depending both on state and time, opening the possibility to include products of state and input.

The numerical computation of periodic solutions in time-periodic non-autonomous systems is a task of greatest interest in engineering application. Periodic solutions are of prime importance for, e.g., nonlinear vibration analysis in structural dynamics [11,12], in acoustics [13] and thermo-acoustics [14], nonlinear oscillation analysis in electronic circuits [15] as well as heart rhythm analysis in cardiology [16]. Therefore, it is important to find and characterize periodic solutions, assess their stability properties and also track these quantities along varying system parameters, a process which is called continuation [17].

Naively, attractive periodic solutions can be found by simply simulating numerically over a long time interval, until transient effects are negligible. However, since this method is very dependent on the initial condition, can only find attractive solutions and is computationally expensive, more sophisticated methods have been developed. There are a multitude of periodic solution solvers available, including finite differences [18], shooting [19], multiple shooting [20], collocation and generalized collocation [21] and harmonic balancing [22]. We highlight two methods:

- The shooting method [19] uses the Newton method and a numerical ODE solver to find initial conditions $\mathbf{x}_0$ such that $\mathbf{x}(t_0 + T) = \mathbf{x}(t_0) = \mathbf{x}_0$, i.e., whose trajectory over a period fulfills the periodicity constraint. Since the monodromy matrix appears in the update step of the Newton method, stability information comes (almost) for free with this method. This monodromy matrix also plays a role in the tangent prediction in an arc-length continuation method.
- The harmonic balance method (HBM) [22,23], in contrast to the other mentioned methods, is a purely frequency-based method. The periodic solution is parameterized globally by a finite set of trigonometric functions, and the coefficients are determined using a residual in the frequency domain. The transfer of nonlinear terms into the frequency domain is non-trivial. In practice, this is achieved either using an alternating frequency and time (AFT) method [24] or by providing a recast of the system in quadratic form [25]. One advantage of the HBM is that it automatically provides a filtering effect on the identified periodic solution.

While stability information about the periodic solution branches comes almost for free in the shooting methods as the monodromy matrix is calculated as a necessary continuation step, stability information and hence information about the bifurcations that occur in a system are not so trivial in the HBM. The Hill matrix provides a frequency-based method to assess the stability of periodic solutions [26,27]. It can be found and constructed easily in the numerical asymptotic method (ANM) [25,28], being a continuation method based on the HBM. Hence, computation of stability through its eigenvalues, which approximate the Floquet exponents, is a viable option.

However, two critical problems make the Hill method often unattractive in practice. On the one hand, from a numerical viewpoint, computing the eigenvalues of the large Hill matrix is computationally expensive and potentially inaccurate [29]. On the other hand, for correct assertion of stability, only a non-trivial subset of these eigenvalues must be considered. This process is known in the literature as *sorting* of Floquet exponent candidates. The determination of this subset is still an active area of research, with the approaches being based on the imaginary parts [23,26,30] and potentially in addition the real parts [31] of the eigenvalues, or alternatively symmetry considerations of the eigenvectors [27,28].

In this work, we propose a different approach for obtaining stability information from the Hill matrix, which circumvents both issues mentioned above. Using the Koopman framework we motivate a novel dynamical systems interpretation of the Hill matrix, which allows to compute an approximation of the monodromy matrix directly (i.e., without computing a large number of eigenvalues and subsequent sorting). The proposed method to find the monodromy matrix from the Hill matrix involves the action of the matrix exponential of the Hill matrix applied to a smaller sparse matrix, followed by a projection to the $n \times n$ monodromy matrix. Finally, the stability of the periodic solution can directly be assessed from the n eigenvalues of the monodromy matrix, known to be the Floquet multipliers.

Parts of the research in this work were presented in a preliminary form at the ENOC2020+2 Conference [32], in particular concerning the proposed choice of the Koopman basis functions in Sect. 3 and the connection to the harmonic balance equations and the Hill matrix. The main (and original) contributions of this work are the novel stability method of Sect. 4, the considerations with respect to the matrix projection and the formal proofs for the theorems in Sect. 3.2.

The paper is structured as follows. Section 2 provides an overview of the notation and gives a theoretical background for the concepts that are central to this work, in particular concerning a selection of topics from Koopman theory as well as frequency-based methods for periodic systems. Section 3 introduces the chosen basis for a Koopman lift on time-periodic systems and states the three central theorems which relate this Koopman lift to the classical frequency-based methods. The proofs of these theorems can be found in Appendix B. Section 4 presents the novel stability method based on the findings from Sect. 3, and the projection to the monodromy matrix as well as the computational effort are discussed. These results are illustrated in Sect. 5 using numerical investigations on two exemplary dynamical systems. Finally, concluding remarks are given in Sect. 6.

2 Theoretical background

In this section, the reader is provided with an overview over the theory in the Koopman framework and the Floquet theory that is necessary for the later parts of the paper.

2.1 Notation and terminology

The frequency-based methods considered in this work rely heavily on being represented as (generalized or classical) Fourier series. Hence, a short overview over Fourier series and the notation that is employed will be given.

Scalar quantities will be represented by Greek or Latin slanted lower case letters. This includes scalar-valued functions as elements of a function space. Vectors in Euclidean space will be represented by Latin bold lowercase letters and matrices will be represented by Latin bold uppercase letters. Tuples of functions are represented by bold font and the distinction to Euclidean space can be drawn from context.

The choice of index is connected to its meaning. The index $l \in \mathbb{N}$ is used for indexing over the states of a system, whereas the index $k \in \mathbb{Z}$ is used for indexing over frequency harmonics. While j may appear as arbitrary auxiliary index, the letter i usually denotes the imaginary number $i^2 = -1$ and is only used for indexing purposes if the indexing context is obvious. The Kronecker delta is denoted by δ_{jl}.

If an inner product $\langle \cdot, \cdot \rangle$ with the usual inner product properties (see, e.g., [33]) is defined on a vector space $\mathcal{F}$, elements $f, g \in \mathcal{F}$ of this space are orthogonal if their inner product is zero. An orthonormal system is a set $\left\{\xi_j\right\}_{j=1}^{D}$, $D \in \mathbb{N} \cup \{\infty\}$ with $\left\langle \xi_i, \xi_j \right\rangle = \delta_{ij}$, and a maximal orthonormal system whose span is a dense subset of the considered vector space is called an orthonormal basis.

By slight abuse of notation, the inner product is extended in this paper to tuples of functions $\mathbf{g} \in \mathcal{F}^l$, $\mathbf{h} \in \mathcal{F}^m$ element-wise via

$$\langle \mathbf{g}, \mathbf{h} \rangle := \begin{pmatrix} \langle g_1, h_1 \rangle & \dots & \langle g_1, h_m \rangle \\ \vdots & \ddots & \vdots \\ \langle g_l, h_1 \rangle & \dots & \langle g_l, h_m \rangle \end{pmatrix}. \tag{1}$$

It can be easily verified that, as a generalization of the conjugate symmetry of the inner product, the relation

$$\langle \mathbf{h}, \mathbf{g} \rangle = \langle \mathbf{g}, \mathbf{h} \rangle^* \tag{2a}$$

holds, where $\mathbf{U}^*$ denotes the conjugate transpose of the matrix $\mathbf{U}$. Moreover, the sesquilinearity of the inner product extends to matrices via

$$\langle \mathbf{U}\mathbf{g}, \mathbf{h} \rangle = \mathbf{U} \langle \mathbf{g}, \mathbf{h} \rangle \tag{2b}$$

$$\langle \mathbf{g}, \mathbf{V}\mathbf{h} \rangle = \langle \mathbf{g}, \mathbf{h} \rangle \mathbf{V}^* \tag{2c}$$

for constant complex matrices $\mathbf{U}, \mathbf{V}$ of appropriate size.

If a space $\mathcal{F}$ has an orthonormal basis $\{\xi_j\}_{j=1}^{D}$ stacked into a tuple $\boldsymbol{\xi}$, it is well-known [33] that elements $g \in \mathcal{F}$ admit a (generalized) Fourier series

$$g = \sum_{j=1}^{D} \langle g, \xi_j \rangle \xi_j = \langle g, \boldsymbol{\xi} \rangle \boldsymbol{\xi}, \tag{3}$$

where the matrix inner product notation was used in the last term. The constant, scalar coefficient $\langle g, \xi_j \rangle$ is called the j-th Fourier coefficient of g. In particular, in the space of trigonometric functions of period T, the functions $u_k : [0\ T) \to \mathbb{C}, t \mapsto \mathrm{e}^{ik\omega t}, k \in \mathbb{Z}$ constitute an orthonormal basis with respect to the inner product

$$\langle f, g \rangle = \int_0^T f(t)\bar{g}(t)\mathrm{d}t. \tag{4}$$

This is the classical Fourier series. For a finite-dimensional subspace spanned by finitely many basis functions $\{u_k\}_{k=-N_\mathbf{u}}^{N_\mathbf{u}}$, a function g can thus be expressed by

$$g = \langle g, \mathbf{u} \rangle \mathbf{u} \tag{5}$$

with the matrix-valued inner product notation introduced earlier.

2.2 Koopman theory overview

A short introduction to the Koopman framework to set the notation and help the reader understand the following sections is given below. It is, however, not intended to give a comprehensive understanding of the current overall state of the art in the field that would be suitable for a wider range of application. For such a more general and in-depth treatment of the considered methodology, the authors rather recommend [10,34].

Consider a non-autonomous time-periodic finite-dimensional dynamical system governed by

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t), \tag{6}$$

where $t \in \mathbb{R}$ is the time, $\mathbf{x}(t) \in \mathcal{X} \subseteq \mathbb{R}^n$ is the state trajectory starting at $\mathbf{x}(t_0) = \mathbf{x}_0$ and $\mathbf{f} : \mathcal{X} \times \mathbb{R} \to \mathcal{X}$ is a smooth vector field which is T-periodic in t. The family of maps $\boldsymbol{\phi}_t(\mathbf{x}_0, t_0) = \mathbf{x}(t)$ characterizes the flow of the system and assigns to each (initial) configuration $(\mathbf{x}_0, t_0)$ the resulting state at time $t \geq t_0$.

The Koopman framework [10] considers output functions $g(\mathbf{x}, t)$, also called observables. Any Banach spaces of functions over the complex or real numbers are permitted in the general Koopman framework. In this work, we consider in particular the space $\mathcal{F}$ of complex-valued functions $g : \mathcal{X} \times \mathbb{R} \to \mathbb{C}$ which are real analytic on $\mathcal{X}$ and T-periodic in the last argument t. Given any function g, it may be of interest how its function values evolve along the trajectories of the system. For instance, in the Lyapunov framework, it is desired that function values of a Lyapunov candidate decrease over time for any starting point. The operator $K^t : \mathcal{F} \to \mathcal{F}; g \mapsto g \circ \boldsymbol{\phi}_t$ performs this shift along the trajectory for arbitrary functions g from the considered function space. The family of all these operators for any t is called the Koopman semigroup of operators. Indeed, the semigroup properties with respect to the time parameter can be verified easily.

For suitable function spaces $\mathcal{F}$, this Koopman semigroup contains all information about the system without explicitly knowing the vector field $\mathbf{f}$ or the flow $\boldsymbol{\phi}_t$. In particular, if $\mathcal{F}$ is chosen such that the identity function id is contained in the vector space, then the flow can be recovered easily by simply evaluating $K^t(\mathrm{id})$. As a trivial counterexample, consider the one-dimensional vector space of constant functions. Any constant function will not change its function value while being evaluated along an arbitrary trajectory of arguments. Therefore, in this case, the Koopman operator semigroup is well defined, albeit trivial. No information about the underlying system is retained. This example shows that an appropriate choice of function space is a crucial part of the Koopman framework.

Under the aforementioned assumptions for the particular function space $\mathcal{F}$ and the vector field $\mathbf{f}$, the Koopman semigroup is continuous with respect to time and there also exists the operator $L : \mathcal{F} \to \mathcal{F}$ with $g \mapsto \dot{g} = \lim_{t \to 0} \frac{K^t g - g}{t}$, mapping an observable g to its total time derivative $\dot{g}$ along the flow with

$$\dot{g}(\mathbf{x}, t) = \frac{\partial g(\mathbf{x}, t)}{\partial \mathbf{x}} \mathbf{f}(\mathbf{x}, t) + \frac{\partial g(\mathbf{x}, t)}{\partial t}. \tag{7}$$

The operator L is called the infinitesimal Koopman generator. Again, for suitable $\mathcal{F}$, this representation alone is a sufficient way to describe the behavior of the

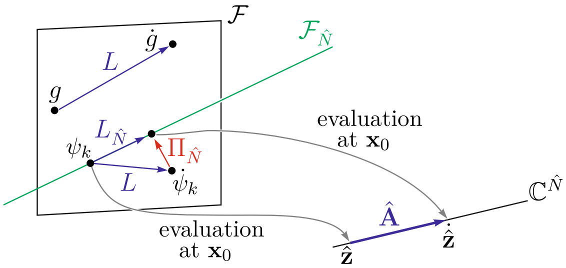


Fig. 1 Schematic drawing of the infinitesimal Koopman generator L, its finite-dimensional approximation $L_{\hat{N}}$ and the Koopman lift $\hat{\mathbf{A}}$

dynamical system. In particular, if id $\in \mathcal{F}$, the vector field $\mathbf{f}$ is easily recovered.

In addition, the infinitesimal Koopman generator and the Koopman semigroup of operators are linear in the argument g, even if the governing differential equation is nonlinear. This comes at the cost of dealing with a mapping on an (infinite-dimensional) function space $\mathcal{F}$ instead of the (finite-dimensional) state space $\mathcal{X}$.

For practical reasons, we will be forced to project the dynamics on a finite-dimensional subspace $\mathcal{F}_{\hat{N}} \subset \mathcal{F}$ spanned by $\hat{N}$ linearly independent basis functions $\{\psi_j\}_{j=1}^{\hat{N}}$. Any projection $\Pi_{\hat{N}} : \mathcal{F} \to \mathcal{F}_{\hat{N}}$ defines a finite-dimensional approximation $\hat{L} : \mathcal{F}_{\hat{N}} \to \mathcal{F}_{\hat{N}}$ of L on $\mathcal{F}_{\hat{N}}$ by $L_{\hat{N}} := \Pi_{\hat{N}} L$. The approximation process and the subsequent approximation error are visualized in Fig. 1. As the subspace $\mathcal{F}_{\hat{N}}$ generally is not closed w.r.t. L, the result of Lg must be projected back onto $\mathcal{F}_{\hat{N}}$, introducing some approximation error. As for $\mathcal{F}$, the choice of the finite space $\mathcal{F}_{\hat{N}}$ and the projection onto it is crucial.

To evaluate a system trajectory in the usual state-space form, the initial condition is fixed first, the state evolution is computed afterward and the output function is computed last. For the Koopman infinitesimal generator (and its approximation on the finite-dimensional function space), this is reversed. First, an observable, or output, is fixed, the observable is propagated along the flow and the initial condition is determined in the last step by evaluating $(Lg)(\mathbf{x})$ for an initial condition $\mathbf{x}$. To arrive back at a linear system representation in the usual state-space form, this behavior must be reversed again. This is achieved by evaluating the action of the (approximate) Koopman infinitesimal generator on the basis functions. Setting

$$\hat{z}_j(0) := \psi_j(\mathbf{x}(0)) \tag{8a}$$

$$\dot{\hat{z}}_j(0) := (L_{\hat{N}} \hat{\psi}_j(\mathbf{x}(0)) \tag{8b}$$

and keeping this dynamics for increasing t, the linear autonomous system

$$\dot{\hat{\mathbf{z}}} = \hat{\mathbf{A}}\hat{\mathbf{z}} \tag{9a}$$

$$\hat{\mathbf{z}}(0) = \hat{\boldsymbol{\Psi}}(\mathbf{x}(0)) := \begin{pmatrix} \psi_1(\mathbf{x}(0)) \\ \vdots \\ \psi_{\hat{N}}(\mathbf{x}(0)) \end{pmatrix} \tag{9b}$$

results. This linear system is called the Koopman lift and describes the dynamics represented in $L_{\hat{N}}$. With the lifted states

$$\hat{\mathbf{z}}(t) \approx \hat{\boldsymbol{\Psi}}(\mathbf{x}(t)) \tag{9c}$$

it is a finite-dimensional linear approximation of the original system dynamics. The Koopman lift matrix $\hat{\mathbf{A}}$ can be derived from the original nonlinear dynamics manually using a column vector $\boldsymbol{\Psi} := \left(\psi_1, \ldots, \psi_{\hat{N}}\right)^{\mathrm{T}}$ of the basis functions of $\mathcal{F}_{\hat{N}}$, computing $\frac{\mathrm{d}\boldsymbol{\Psi}}{\mathrm{d}t}$ and identifying terms linear in elements of $\boldsymbol{\Psi}$ after applying the projection $\Pi_{\hat{N}}$ from $\mathcal{F}$ onto $\mathcal{F}_{\hat{N}}$. If this projection is orthonormal and $\boldsymbol{\Psi}$ is an orthonormal system, then the matrix entry $\hat{\mathbf{A}}_{i,j}$ at i-th row and j-th column is given by $\hat{\mathbf{A}}_{i,j} = \left\langle \frac{\mathrm{d}\psi_i}{\mathrm{d}t}, \psi_j \right\rangle$, where $\langle ., . \rangle$ denotes the corresponding inner product.

Depending on the basis structure chosen, various popular embedding techniques emerge as a Koopman lift for specific system classes. For instance, if a monomial basis is chosen for an autonomous system, the Carleman linearization [35] results as Koopman lift. For smooth, polynomial systems, this is often the first choice of basis dictionary [36]. Alternatively, delay coordinates are often employed in Koopman-based applications [2]. Based on the Takens embedding theorem, this can capture weakly nonlinear dynamics [37]. For periodic systems, the Fourier embedding [38] has been known, although its properties have mainly been analyzed in the frequency domain.

2.3 Harmonic balance method

Consider the non-autonomous time-periodic finite-dimensional dynamical system (6) as above. Often one is interested in finding a T-periodic solution, i.e., a solution to the dynamical system (6) which fulfills

$\mathbf{x}(t+T) = \mathbf{x}(t)$ for all $t \geq 0$. This constitutes a boundary value problem (BVP) and there are methods to solve this type of BVP in the time domain and in the frequency domain. Shooting, multiple shooting and collocation methods all rely on an interplay between time-integration (or finite differencing) of the ODE and solving nonlinear functions for periodicity and continuity constraints [21].

In contrast, the HBM is a frequency-based method. Under suitable smoothness assumptions, the periodic solution has a convergent Fourier series. Hence the periodic solution can be approximated by its Fourier expansion up to order N_{HBM} with unknown parameters via

$$\mathbf{x}_p(t) = \sum_{k=-N_{\text{HBM}}}^{N_{\text{HBM}}} \mathbf{p}_k \mathrm{e}^{ik\omega t} \tag{10}$$

with $\omega = \frac{2\pi}{T}$, $\mathbf{u}(t) = (\mathrm{e}^{-iN_{\text{HBM}}\omega t}, \ldots, \mathrm{e}^{iN_{\text{HBM}}\omega t})$ being a vector of Fourier base functions and $\{\mathbf{p}_{-N_{\text{HBM}}}, \ldots, \mathbf{p}_{N_{\text{HBM}}}\}$ gathering the corresponding (unknown) coefficients. These coefficients $\mathbf{p}_k$ are then determined by substituting this ansatz into the system equation (6). The comparison of coefficients for the Fourier expansions of $\frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t}$ from the definition (10) and $\mathbf{f}(\mathbf{x}_p, t)$ for every order up to N_{HBM} transforms the BVP into a system of $n(2N_{\text{HBM}}+1)$ algebraic equations. Existence and convergence of these HBM approximations has been shown [22]. While the left-hand side of the equation as well as linear terms in $\mathbf{f}$ are easy to handle, the frequency component of the nonlinear terms can usually not be expressed in closed form. The individual equations for each order are thus usually determined and simultaneously solved using the fast Fourier transform with an alternating frequency and time (AFT) method [24]. The equations for each order can also be isolated by projecting onto the corresponding basis function from the collection in $\mathbf{u}$ through the classical inner product (4). Hence, the HBM approximates a periodic solution by solving the $n(2N_{\text{HBM}}+1)$ algebraic equations collected in

$$\left\langle \frac{\mathrm{d}\mathbf{x}_p}{\mathrm{d}t}, \mathbf{u} \right\rangle = \langle \mathbf{f}(\mathbf{x}_p(\cdot), \cdot), \mathbf{u} \rangle . \tag{11}$$

With this notation, the numerically cumbersome task of calculating the Fourier coefficients of the nonlinear components of $\mathbf{f}$ is hidden in the definition of the inner product.

2.4 Floquet theory: stability of periodic solutions

When a periodic orbit $\mathbf{x}_p$ is found (via HBM or by other means), the next interesting question is that of its stability properties; that is, whether trajectories that start sufficiently close to the periodic orbit will approach it, stay in a vicinity of it or tend away from it with increasing time. To evaluate the stability properties, the dynamics of a perturbation $\mathbf{y} = \mathbf{x} - \mathbf{x}_p$ from the periodic solution is considered. Substitution of this definition into the original system dynamics yields

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{x}_p + \mathbf{y}, t) - \dot{\mathbf{x}}_p := \mathbf{J}(t)\mathbf{y} + \mathcal{O}(\|\mathbf{y}\|^2), \tag{12}$$

where $\mathbf{J}(t) = \left.\frac{\partial \mathbf{f}}{\partial \mathbf{x}}\right|_{\mathbf{x}_p(t),t}$ is the Jacobian of the system evaluated along the periodic solution. The approximate linear time-varying (LTV) system

$$\dot{\mathbf{y}}(t) = \mathbf{J}(t)\mathbf{y}(t) \tag{13a}$$

$$\mathbf{y}(0) = \mathbf{x}(0) - \mathbf{x}_p(0) \tag{13b}$$

has an equilibrium at zero, which corresponds to the periodic orbit of the original system, and the stability analysis of the periodic orbit reduces (in the hyperbolic case) to the stability analysis of this equilibrium. This will be the convention for the remainder of this paper, unless stated otherwise.

The fundamental solution matrix $\mathbf{\Phi}(t)$ is the solution to the variational equation

$$\dot{\mathbf{\Phi}}(t) = \mathbf{J}(t)\mathbf{\Phi}(t)\,; \qquad \mathbf{\Phi}(0) = \mathbf{I} \tag{14}$$

and any state can be obtained via $\mathbf{y}(t) = \mathbf{\Phi}(t)\mathbf{y}_0$. In particular, the fundamental solution matrix $\mathbf{\Phi}(T) =: \mathbf{\Phi}_T$ evaluated after one period is called the monodromy matrix of the system and its eigenvalues $\{\lambda_l\}_{l=1}^n$ are called Floquet multipliers [39]. The Poincaré map $\mathbf{y}_{k+1} = \mathbf{\Phi}_T \mathbf{y}_k$ provides snapshots for the evolution of the perturbation $\mathbf{y}$, spaced at a time distance of T, i.e., $\mathbf{y}_k = \mathbf{y}(kT)$. For the long-term behavior, it is sufficient to consider the evolution of these snapshots. Therefore, stability analysis of the periodic solution reduces to stability analysis of the Poincaré map. Hence, if all Floquet multipliers are of magnitude strictly less than one, the equilibrium of the perturbed LTV system and thus the periodic solution of the original system are asymptotically stable; if at least one eigenvalue has a magnitude strictly larger than one, they are unstable. If there exist Floquet multipliers with magnitude equal to one, but none with a magnitude larger than one, the equilibrium is non-hyperbolic and further investigation is necessary to give conclusive statements about stability of the originally considered periodic solution.

Alternatively to the Floquet multipliers, the stability properties of a time-periodic linear system can be characterized by the Floquet exponents. In the linearized perturbed system (13a), if the matrix $\mathbf{\Phi}_T$ is diagonalizable, there exist n solutions $\mathbf{y}_l(t) = \mathbf{p}_l(t)\mathrm{e}^{\alpha_l t}$ which form a basis of the solution space, where each function $\mathbf{p}_l$ is T-periodic [39]. Hence, stability is characterized by the real parts of the Floquet exponents $\{\alpha_l\}_{l=1}^{n}$. If at least one Floquet exponent lies in the open right half plane, i.e., if at least one real part is larger than zero, the equilibrium is unstable. The Floquet multipliers can be determined by substituting $t = T$ in the Floquet solution and it follows that

$$\lambda_l = \mathrm{e}^{\alpha_l T}, \quad l = 1, \ldots, n. \tag{15}$$

In contrast to the Floquet multipliers, the Floquet exponents are not uniquely defined. It is easy to see that if the pair $(\mathbf{p}_l(t), \alpha_l)$ generates a solution $\mathbf{y}_l(t)$, the same solution is generated by $(\tilde{\mathbf{p}}_l(t), \tilde{\alpha}_l) = (\mathbf{p}_l(t)\mathrm{e}^{-ik\omega t}, \alpha_l + ik\omega)$ with $k \in \mathbb{Z}$. Hence, in total, there are infinitely many valid Floquet exponents, which can be categorized into n distinct groups. All elements of one group have the same real part and differ in the imaginary part by multiples of $i\omega$. As stability is determined by the real part only, it is sufficient for stability analysis to know any one element from each of the n groups. All elements of one group map to the same Floquet multiplier.

When a periodic orbit is determined using the purely time-domain-based shooting method, the monodromy matrix usually is a direct byproduct of the continuation method [17]. In this case, the numerically obtained monodromy matrix can be evaluated directly to obtain the Floquet multipliers and their stability information.

When the HBM is computed in the standard way, however, stability information about the identified limit cycle is unclear without further investigation. The Hill method [22,40] offers a frequency-domain-based way to approximate the Floquet exponents of the linearized perturbation equation.

The Floquet exponents are eigenvalues of the infinite Hill matrix $\mathbf{H}_\infty$ [30], which is constructed from the Fourier coefficients of the periodic system matrix $\mathbf{J}(t) = \sum_{k=-\infty}^{\infty} \mathbf{J}_k \mathrm{e}^{i\omega k t}$ and reads as

$$\mathbf{H}_\infty = \begin{pmatrix} \ddots & \vdots & \vdots & \vdots & \iddots \\ \ldots & \mathbf{J}_0 + i\omega\mathbf{I} & \mathbf{J}_{-1} & \mathbf{J}_{-2} & \ldots \\ \ldots & \mathbf{J}_1 & \mathbf{J}_0 & \mathbf{J}_{-1} & \ldots \\ \ldots & \mathbf{J}_2 & \mathbf{J}_1 & \mathbf{J}_0 - i\omega\mathbf{I} & \ldots \\ \iddots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \tag{16}$$

Introducing a vector $\mathbf{v}_\infty$ of appropriate (infinite) length, the eigenproblem

$$\mathbf{H}_\infty \mathbf{v}_\infty = \tilde{\alpha}\mathbf{v}_\infty \tag{17}$$

can be formulated. This infinite-dimensional problem has infinitely many discrete eigenvalues $\tilde{\alpha}$, which solve (17). They correspond identically to the Floquet exponents $\tilde{\alpha}$ for all $k \in \mathbb{Z}$ as introduced above and can be sorted into n groups, where the entries of each group differ by multiples of $i\omega$ [30]. However, in practice, only the eigenvalues of a finite-dimensional matrix approximation of $\mathbf{H}_\infty$ can be computed numerically. The matrix

$$\mathbf{H} = \begin{pmatrix} \mathbf{J}_0 + iN_{\mathbf{u}}\omega\mathbf{I} & \ldots & \mathbf{J}_{-2N_{\mathbf{u}}} \\ \vdots & \ddots & \vdots \\ \mathbf{J}_{2N_{\mathbf{u}}} & \ldots & \mathbf{J}_0 - iN_{\mathbf{u}}\omega\mathbf{I} \end{pmatrix} \tag{18}$$

of size $n(2N_{\mathbf{u}} + 1) \times n(2N_{\mathbf{u}} + 1)$ consists of the $n(2N_{\mathbf{u}} + 1)$ most centered rows and columns of $\mathbf{H}_\infty$ and approximates the original infinite-dimensional Hill matrix. In the absence of truncation error, the eigenvalues of $\mathbf{H}$ would be a subset of the eigenvalues of $\mathbf{H}_\infty$, i.e., the Floquet exponents. Due to the inevitable error which generally comes with truncation, however, this does not quite hold. The N eigenvalues of $\mathbf{H}$, which do not identically coincide with Floquet exponents, will be called Floquet exponent candidates below.

The matrix $\mathbf{H}$ has a block Toeplitz structure except for the middle diagonal, and, for sufficiently large $N_{\mathbf{u}}$, the bands near the diagonal dominate as the Fourier coefficients of $\mathbf{J}$ tend to zero. Loosely speaking, some eigenvalues affiliated most with the central rows of $\mathbf{H}$ are less impacted by the truncation and provide a better approximation to the Floquet exponents than others [28]. Up until the change of the last century, this property was neglected and stability of the periodic solution was asserted based on the real parts of all Floquet exponent candidates [40]. This naive Hill method without any additional steps would often assert instability for stable solutions due to spurious Floquet exponent candidates without physical meaning, giving it a reputation of being inaccurate [22,29].

However, the accuracy of the Hill method can be improved significantly if only a subset of Floquet exponent candidates is considered, instead of all of them. Hence, the search for a selection criterion which determines the best approximation to the Floquet exponents from the Floquet candidates has received much attention in the literature [23,26,28,30].

For sufficiently large $N_{\mathbf{u}}$, it is proven that the candidates with minimal imaginary part in modulus converge to the true Floquet exponents [26]. Since the convergence may only occur for very large truncation orders, an addition to this method was very recently proposed [31]. This modified method first sorts the Floquet exponent candidates based on their real parts, before applying the imaginary part criterion to the most highly populated groups. However, in this real-part-based method, it is unclear how to proceed if two or more true Floquet exponents have the same real parts, and thus not enough groups are available.

An alternative criterion selects those candidates whose eigenvectors are most symmetric [27,28], as they should correspond most to the middle rows of the matrix **H**. This symmetry is computed based on a weighted mean. Even though there currently is no formal convergence proof for this symmetry-based sorting method, some numerical results indicate faster convergence than with the aforementioned eigenvalue criterion [12], while other results do not support this claim [31].

For all these criteria, there are currently no methods to efficiently and accurately compute only those eigenvalues which fit the criterion. Rather, all eigenpairs of the large matrix have to be computed first and then most of them are discarded. As the cost of solving an eigenvalue problem of a $N \times N$ matrix is of the order $\mathcal{O}(N^3)$, the computational cost of the approach is usually dominated by determining the eigendecomposition of a large matrix [29]. In addition, it is well known that the accuracy of the computed eigenvalues is not too high if all eigenvalues of a sparse matrix are sought. Sparsity of **H** cannot be reasonably exploited to decrease the computational cost of solving the complete eigenproblem [41].

3 A Koopman dictionary for time-periodic systems

For smooth autonomous systems in the Koopman framework, it is customary to choose as basis functions the so-called Carleman basis [2,42], i.e., a finite set of monomials $\psi_{\boldsymbol{\beta}}(\mathbf{x}) = \mathbf{x}^{\boldsymbol{\beta}}$, where $\boldsymbol{\beta} \in \mathbb{N}^n$ is a multi-index and standard multi-index calculation rules (see Appendix A) apply. As time-periodic functions are considered here, we propose in this paper to include as basis functions combinations of monomial terms as well as Fourier terms of the base frequency, i.e., basis functions of the form $\psi_{\boldsymbol{\beta},k} := \mathbf{x}^{\boldsymbol{\beta}} \mathrm{e}^{ik\omega t}$, where $\omega = \frac{2\pi}{T}$. The functions $\left\{\psi_{\boldsymbol{\beta},k} | k \in \mathbb{Z}, \boldsymbol{\beta} \in \mathbb{N}_0^n\right\}$ are an orthonormal system within the initially considered vector space $\mathcal{F}$ w.r.t. the inner product

$$\langle g, h \rangle := \int_0^T \left(\frac{1}{(\boldsymbol{\beta}!)^2} \sum_{\boldsymbol{\beta} \in \mathbb{N}^n} \left. \frac{\partial^{\boldsymbol{\beta}} g}{\partial \mathbf{x}^{\boldsymbol{\beta}}} \right|_{\mathbf{0},t} \left. \frac{\partial^{\boldsymbol{\beta}} \bar{h}}{\partial \mathbf{x}^{\boldsymbol{\beta}}} \right|_{\mathbf{0},t} \right) \mathrm{d}t. \tag{19}$$

This inner product contains the standard inner product for Fourier series, and the derivatives serve as an inner product for the monomials. The inner product properties can be readily verified.

Let $N_{\mathbf{z}}, N_{\mathbf{u}} \in \mathbb{N}$ be integers which describe the assumed maximum polynomial and frequency order, respectively. There is a set $\mathcal{B} = \{\boldsymbol{\beta}_j\}_{j=1}^{N_{\boldsymbol{\beta}}}$ collecting all multi-indices with $1 \leq \|\boldsymbol{\beta}\| \leq N_{\mathbf{z}}$. These are all multi-indices that create monomials $\mathbf{x}^{\boldsymbol{\beta}}$ of degree $N_{\mathbf{z}}$ and less. By (A4), it holds that $N_{\boldsymbol{\beta}} = \binom{N_{\mathbf{z}}+n}{n} - 1$. For the sake of brevity, define $N = N_{\boldsymbol{\beta}}(2N_{\mathbf{u}} + 1)$ and $\hat{N} = N + (2N_{\mathbf{u}} + 1)$. The set

$$\left\{\psi_{\boldsymbol{\beta},k} | \boldsymbol{\beta} \in \mathcal{B} \cup \{\mathbf{0}\}, |k| \leq N_{\mathbf{u}}\right\} \tag{20}$$

of orthogonal basis functions spans a specific finite-dimensional subspace $\mathcal{F}_{\hat{N}} \subset \mathcal{F}$. These basis functions are the monomials up to degree $N_{\mathbf{z}}$, multiplied to the Fourier base functions up to frequency order $N_{\mathbf{u}}$. With $k = 0$ and $\|\boldsymbol{\beta}\| = 1$, the set (20) includes the identity function for each state. Moreover, since the multi-index $\boldsymbol{\beta} = \mathbf{0}$ is permitted in the set (20), the purely time-dependent classical Fourier base functions $\mathrm{e}^{ik\omega t}$ are represented. These basis functions are collected into two vectors. The vector

$$\mathbf{u}(t)^{\mathrm{T}} := (\mathrm{e}^{-i\omega N_{\mathbf{u}} t}, \dots, \mathrm{e}^{0}, \dots, \mathrm{e}^{i\omega N_{\mathbf{u}} t}) \tag{21}$$

collects all basis functions that are not dependent on the state, but only on time, such that the evolution of **u** is not a product of the system dynamics, but is known a priori. All other state-dependent basis functions are collected into the vector

$$\boldsymbol{\Psi}_{\mathbf{z}}(\mathbf{x}, t) := \begin{pmatrix} \mathbf{x}^{\beta_1} \mathrm{e}^{-i\omega N_{\mathbf{u}} t} \\ \vdots \\ \mathbf{x}^{\beta_1} \mathrm{e}^{0} \\ \vdots \\ \mathbf{x}^{\beta_1} \mathrm{e}^{i\omega N_{\mathbf{u}} t} \\ \mathbf{x}^{\beta_2} \mathrm{e}^{-i\omega N_{\mathbf{u}} t} \\ \vdots \\ \mathbf{x}^{\beta_{N_\beta}} \mathrm{e}^{i\omega N_{\mathbf{u}} t} \end{pmatrix}. \tag{22}$$

Here, the basis functions are ordered by monomial exponent first, and then all frequency base functions for the same monomial are grouped together in ascending order. It is notable that any other orderings are also applicable and the matrices can be transformed into each other by similarity transforms. The ordering (22) has been chosen purely for convenience in later argumentation.

If the full basis as introduced above is considered, the state **x** itself is included in the basis. Hence, there is a selector matrix $\mathbf{C}_{\mathbf{z}} \in \mathbb{R}^{n \times N}$ containing select rows of the identity matrix with $\mathbf{C}_{\mathbf{z}} \boldsymbol{\Psi}_{\mathbf{z}}(\mathbf{x}, t) = \mathbf{x}$. There exist other options to recover **x** from $\boldsymbol{\Psi}_{\mathbf{z}}(\mathbf{x}, t)$. For instance, monomials of higher (uneven) order can be used by considering the corresponding root. Also, the matrix $\mathbf{C}_{\mathbf{z}}$ can be allowed to be time-dependent. This second option and its implications will be investigated in more detail in Sect. 4.3. Both vectors $\boldsymbol{\Psi}_{\mathbf{z}}$ and **u** are collected into one vector $\boldsymbol{\Psi} := (\boldsymbol{\Psi}_{\mathbf{z}}^{\mathrm{T}}, \mathbf{u}^{\mathrm{T}})^{\mathrm{T}}$ for convenience.

As introduced in Sect. 2.2, the Koopman lift describes the evolution of the basis functions under the approximate infinitesimal Koopman generator, and it is determined by expressing the time derivative along the flow of all basis functions in the finite-dimensional basis, projecting them back again to the subspace. Using the projection defined by the inner product (19), this gives

$$\dot{\boldsymbol{\Psi}} = \left\langle \dot{\boldsymbol{\Psi}}, \boldsymbol{\Psi} \right\rangle \boldsymbol{\Psi} + \mathbf{r} \,, \tag{23}$$

where $\mathbf{r} \in \mathcal{F}^{\hat{N}}$ is the remainder that is orthogonal to the finite-dimensional subspace and will be projected out. Substituting (23) into the definition (9) for the Koopman lift and separating the two vectors of basis functions yields

$$\dot{\hat{\mathbf{z}}} := \begin{pmatrix} \dot{\mathbf{z}} \\ \dot{\mathbf{u}} \end{pmatrix} = \left\langle \dot{\boldsymbol{\Psi}}, \boldsymbol{\Psi} \right\rangle \hat{\mathbf{z}} \tag{24a}$$

$$= \begin{pmatrix} \left\langle \dot{\boldsymbol{\Psi}}_{\mathbf{z}}, \boldsymbol{\Psi}_{\mathbf{z}} \right\rangle & \left\langle \dot{\boldsymbol{\Psi}}_{\mathbf{z}}, \mathbf{u} \right\rangle \\ \left\langle \dot{\mathbf{u}}, \boldsymbol{\Psi}_{\mathbf{z}} \right\rangle & \left\langle \dot{\mathbf{u}}, \mathbf{u} \right\rangle \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ \mathbf{u} \end{pmatrix} \tag{24b}$$

$$=: \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{0} & * \end{pmatrix} \begin{pmatrix} \mathbf{z} \\ \mathbf{u} \end{pmatrix} \,, \tag{24c}$$

where $\mathbf{A} \in \mathbb{C}^{N \times N}$ and $\mathbf{B} \in \mathbb{C}^{N \times (2N_{\mathbf{u}}+1)}$ are constant coefficient matrices. The lower rows of the large matrix in (24b) determine the dynamics of **u**. However, since **u** is purely time-dependent and its time evolution is known a priori, the lower rows of (24b) are superfluous and the original nonlinear system is approximated by the LTI system

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\mathbf{u} \tag{25a}$$

$$\mathbf{z}(0) = \boldsymbol{\Psi}_{\mathbf{z}}(\mathbf{x}_0, 0) \tag{25b}$$

for the very specific input **u** as in (21). The lifted state vector $\mathbf{z}(t)$ is an approximation to the state-dependent basis functions $\boldsymbol{\Psi}_{\mathbf{z}}(\mathbf{x}(t), t)$ evaluated along the flow of the original system.

This is only an approximation, however, due to the projection onto the finite-dimensional function space. In particular, for $t \geq 0$, the lifted state **z** will cease to adhere to the constraints posed by $\boldsymbol{\Psi}_{\mathbf{z}}$, meaning that there may not exist any vector $\tilde{\mathbf{x}} \in \mathbb{R}^n$ which fulfills $\mathbf{z}(t) = \boldsymbol{\Psi}_{\mathbf{z}}(\tilde{\mathbf{x}}, t)$.

Many results of this work tie in to classical results that are obtained from Jacobian linearization of the state space. Therefore, the maximum monomial order in the basis will often be restricted to $N_{\mathbf{z}} = 1$. In this case, the vector of basis functions will be denoted by $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}$ of size $n(2N_{\mathbf{u}} + 1)$ and for the sake of brevity, these functions will be called linear basis functions, even though the influence of time is still decidedly nonlinear. To make the ordering unique, these linear basis functions are ordered in an ascending order, first by state and then by frequency such that the resulting vector reads as

$$\mathbf{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{x},t) = \begin{pmatrix} x_1 \mathrm{e}^{-iN_{\mathbf{u}}\omega t} \\ \vdots \\ x_1 \mathrm{e}^{iN_{\mathbf{u}}\omega t} \\ x_2 \mathrm{e}^{-iN_{\mathbf{u}}\omega t} \\ \vdots \\ x_n \mathrm{e}^{iN_{\mathbf{u}}\omega t} \end{pmatrix}. \tag{26}$$

3.1 Koopman lift on the perturbed system

While a nonlinear dynamical system may have an arbitrary number of attractors of any kind (equilibrium, periodic, quasi-periodic, chaotic) located anywhere in the state space, the finite-dimensional linear Koopman lift (25a) has a very limited range of possible attractors, i.e., at most one single globally attractive solution, which is an equilibrium point. This means that the system (25a) will most likely not be able to describe the system dynamics of a nonlinear time-periodic system globally. For time-periodic systems, periodic solutions are expected. In this section, the Koopman lift of the perturbed system around such a periodic solution will be regarded to assess its local behavior.

As in the standard HBM, a periodic solution is approximated by its Fourier expansion (10) up to order N_{HBM} with unknown Fourier coefficients. The perturbed system is then given by $\mathbf{y}(t) = \mathbf{x}(t) - \mathbf{x}_p(t)$, with dynamics

$$\dot{\mathbf{y}}(t) := \tilde{\mathbf{f}}(\mathbf{y}(t), t) \tag{27a}$$

as in (12). This allows the Koopman lift to be performed on the perturbed system, i.e., on functions of the state $\mathbf{y}$ evaluated along the flow $\tilde{\mathbf{f}}$. Now, the origin of the lifted state does correspond to the origin equilibrium of the perturbed dynamics, or a periodic solution of the original dynamics. This also means that the Koopman lift matrices $\mathbf{A}$ and $\mathbf{B}$ now depend on the Fourier coefficients $\mathbf{p}$ of the periodic solution around which we are linearizing. The following sections shed light on the properties of the matrices $\mathbf{A}(\mathbf{p})$ and $\mathbf{B}(\mathbf{p})$.

3.2 Koopman-based harmonic balance and Hill equations

One of our key findings is that the $\mathbf{B}(\mathbf{p})$ matrix of the Koopman lift contains information about the parameterization of the periodic solution that is also encoded in the HBM. This is made more precise in the theorems given in this section.

Let $\mathbf{x}_p = \sum_k \mathbf{p}_k \mathrm{e}^{ik\omega t}$ be a real-valued periodic solution candidate of the time-periodic system (6). The harmonic residual of this candidate is given by

$$\mathbf{r}(t) = \mathbf{f}(\mathbf{x}_p(t), t) - \dot{\mathbf{x}}_p(t). \tag{28}$$

Since $\mathbf{x}_p$ as well as $\mathbf{f}$ are periodic, the residual is periodic as well, and therefore it has a Fourier series

$$\mathbf{r}(t) = \sum_k \mathbf{r}_k \mathrm{e}^{ik\omega t}. \tag{29}$$

The HBM of order M returns solution candidates for which $\mathbf{r}_k = 0$ for all $|k| \leq M$. Since $\mathbf{r}$ is real-valued, it also holds that $\mathbf{r}_{-k} = \bar{\mathbf{r}}_k$, where $\mathbf{r}_k = [r_{k,1}, \ldots, r_{k,n}]^{\mathrm{T}}$ is a vector with n complex-valued entries.

This definition of the residual allows to concisely relate the HBM to the Koopman lift.

Theorem 1 *Let $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{p})\mathbf{z} + \mathbf{B}(\mathbf{p})\mathbf{u}$ be the lifted dynamics of frequency order N_{HBM} of system* (6) *around an unknown periodic ansatz of the form* (10). *The N_{HBM}-th order HBM equations* (11), *i.e., $\mathbf{r}_k = 0$, $|k| \leq N_{\mathrm{HBM}}$, are given by $\mathbf{C}_{\mathbf{z}}\mathbf{B}(\mathbf{p}) = \mathbf{0}$, where $\mathbf{C}_{\mathbf{z}}$ is the constant selection matrix that fulfills $\mathbf{y} = \mathbf{C}_{\mathbf{z}}\mathbf{\Psi}_{\mathbf{z}}(\mathbf{y}, t)$ for all t.*

The formal proof of this theorem is given in Appendix B.1. From an intuitive point of view, this condition is not surprising. For a periodic solution, the lifted dynamics has an equilibrium at zero, meaning that if $\mathbf{y} = \mathbf{0}$ it should also hold that $\dot{\mathbf{y}} = \mathbf{0}$. Since the lifted dynamics approximates $\mathbf{z}(t) \approx \mathbf{\Psi}_{\mathbf{z}}(\mathbf{y}, t)$ and this approximation holds identically at the initial point (25b), it should hold that $\dot{\mathbf{y}} \approx \mathbf{C}_{\mathbf{z}}\dot{\mathbf{z}}$ is zero if $\mathbf{y} = \mathbf{0}$. This equation can be evaluated for an arbitrary initial point on the periodic ansatz via a time-shift. With $\mathbf{y}(0) = \mathbf{0}$ and $\mathbf{z}(0) = \mathbf{\Psi}_{\mathbf{z}(\mathbf{y}(0),0)} = \mathbf{0}$, it turns out that the only remaining summand in the approximate $\mathbf{y}$-dynamics is $\mathbf{C}_{\mathbf{z}}\mathbf{B}(\mathbf{p})$.

If only basis functions that are linear in the state ($N_{\mathbf{z}} = 1$) are considered, the above result still holds.

Moreover, we can state Theorem 2 about all entries of $\mathbf{B}$ and not just specific rows.

Theorem 2 *Let* $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{p})\mathbf{z} + \mathbf{B}(\mathbf{p})\mathbf{u}$ *be the lifted dynamics of system* (6) *with linear basis functions* $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}$ *of frequency order* $N_{\mathbf{u}}$ *that are sorted as in* (22), *evaluated for the perturbed system around an unknown periodic ansatz of the form* (10) *up to frequency order at least* $N_{\mathrm{HBM}} = 2N_{\mathbf{u}}$. *Then, the matrix* $\mathbf{B}(\mathbf{p}) \in \mathbb{C}^{n(2N_{\mathbf{u}}+1)\times(2N_{\mathbf{u}}+1)}$ *consists of n stacked Toeplitz matrices. The* l*-th Toeplitz matrix* $\mathbf{B}_l$ *contains as entries (ignoring duplicates) precisely the* $4N_{\mathbf{u}} + 1$ *residuals* $r_{k,l}(\mathbf{p})$, $|k| \leq 2N_{\mathbf{u}}$ *that follow from the HBM w.r.t the* l*-th state.*

If $\mathbf{B}(\mathbf{p}) = \mathbf{0}$, *then all these residuals of the HBM vanish. Conversely, if* $\mathbf{p}$ *solves the HBM equations* $\mathbf{r}_k(\mathbf{p}) = \mathbf{0}$, $|k| \leq 2N_{\mathbf{u}}$, *then it holds that* $\mathbf{B}(\mathbf{p}) = \mathbf{0}$.

The formal proof of this theorem is given in Appendix B.1.

In addition to the $\mathbf{B}$ matrix, the $\mathbf{A}$ matrix of the Koopman lift also holds frequency information about stability of the periodic solution. This is summarized in the following theorem.

Theorem 3 *Let* $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{p})\mathbf{z} + \mathbf{B}(\mathbf{p})\mathbf{u}$ *be the lifted dynamics around a periodic solution of system* (6) *with linear basis functions* $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}$ *of frequency order* $N_{\mathbf{u}}$ *that are ordered as in* (22). *Then the Hill matrix* $\mathbf{H}$, *truncated to frequency order* $N_{\mathbf{u}}$, *for the periodic solution parameterized by* $\mathbf{p}$ *results from the matrix* $\mathbf{A}(\mathbf{p})$ *by the similarity transform* $\mathbf{H} = \mathbf{U}\mathbf{A}(\mathbf{p})\mathbf{U}^{\mathrm{T}}$, *where* $\mathbf{U}$ *is an orthogonal permutation matrix that satisfies* $\mathbf{U}\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}} = (\mathbf{y}^{\mathrm{T}}\mathrm{e}^{iN_{\mathbf{u}}\omega t}, \ldots, \mathbf{y}^{\mathrm{T}}\mathrm{e}^{-iN_{\mathbf{u}}\omega t})^{\mathrm{T}}$.

The formal proof of this theorem, again based on explicit evaluation of the inner product in the Koopman lift, is given in Appendix B.2.

With the three above theorems, qualitative insight about the accuracy of the presented Koopman lift can be gained. Locally (in the vicinity of a periodic solution), the lifted system contains the same dynamical information that is encapsulated in the Hill matrix of the same frequency order. Convergence results about the HBM [22] and the Hill method [26] can thus be related to the accuracy of the Koopman lift. Within the applied Koopman community, such a link is quite unusual. For many applications, convergence results for the finite-dimensional Koopman lift do not exist at all. The convergence results that do exist (e.g., [43,44]) state that, under special conditions, a given error tolerance can be reached using a large number of specific basis functions, without giving an explicit upper bound. Hence, for time-periodic systems of the form (6), the proposed class of basis functions is a favorable choice due to its added connection with the Hill matrix, which indicates that the Koopman lift retains valuable stability information.

4 Sorting-free stability method

The Koopman lift with Theorems 2 and 3 gives a dynamical systems interpretation for the Hill matrix, allowing for a novel stability method based on the Hill matrix. This is demonstrated in the following section.

4.1 Approximating the monodromy matrix

Consider the Koopman lift as in Theorems 2 and 3, i.e., consider the linear basis $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}$ evaluated for the perturbed system around a periodic solution which is determined up to a frequency order of $2N_{\mathbf{u}}$. From Theorem 2, we know that $\mathbf{B} = \mathbf{0}$ and from Theorem 3, that $\mathbf{A} = \mathbf{U}^{\mathrm{T}}\mathbf{H}\mathbf{U}$. The linear dynamical system $\dot{\mathbf{z}} = \mathbf{A}\mathbf{z}+\mathbf{B}\mathbf{u}$ resulting from the Koopman lift therefore reduces to

$$\dot{\mathbf{z}}(t) = \mathbf{A}\mathbf{z}(t) \tag{30a}$$

$$\mathbf{y}(t) \approx \mathbf{z}_{\mathbf{y}}(t) = \mathbf{C}(t)\mathbf{z}(t) \tag{30b}$$

$$\mathbf{z}(0) = \boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{y}(0), 0)\ , \tag{30c}$$

where $\mathbf{C}(t) \in \mathbb{C}^{n\times N}$ is a possibly time-dependent projection matrix that satisfies $\mathbf{C}(t)\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{y}, t) = \mathbf{y}$ for all $t \in [0, T)$. There is a $nN_{\mathbf{u}}$-parameter family of choices for $\mathbf{C}$ if it is allowed to be time-dependent, which will be investigated further in Sect. 4.3. However, the naive choice would be to pick the entries in (26) that correspond to frequency zero. In this case, the matrix $\mathbf{C}$ is constant and given by

$$\mathbf{C} = \mathbf{I}_{n\times n} \otimes \begin{pmatrix} 0 \ldots 0\ 1\ 0 \ldots 0 \end{pmatrix}\ , \tag{31}$$

where $\mathbf{I}_{n\times n}$ is the $n \times n$ identity matrix, ⊗ denotes the Kronecker product and the second matrix in (31) is a row vector $\in \mathbb{R}^{2N_{\mathbf{u}}+1}$ with zeros everywhere except for the middle column.

Since for $t = 0$ all exponential terms are 1 and vanish in (26), the vector $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{y}, 0)$ can be expressed as a matrix product

$$\mathbf{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{y}, 0) = \mathbf{W}\mathbf{y} \tag{32}$$

for all $\mathbf{y}$. The matrix $\mathbf{W} \in \mathbb{R}^{N\times n}$ consists of (repeated) rows of the identity matrix. More specifically, it can be expressed as

$$\mathbf{W} = \mathbf{I}_{n\times n} \otimes \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} , \tag{33}$$

where the second matrix in (33) is a column vector $\in \mathbb{R}^{2N_\mathbf{u}+1}$ filled with ones. As the system (30a)–(30c) is a linear time-invariant (LTI) system (except for the possibly time-dependent matrix $\mathbf{C}$), its closed form solution can be explicitly computed as

$$\mathbf{y}(t) \approx \mathbf{z}_\mathbf{y}(t) = \mathbf{C}(t)\mathrm{e}^{\mathbf{A}t}\mathbf{z}(0) \tag{34a}$$

$$= \mathbf{C}(t)\mathrm{e}^{\mathbf{A}t}\mathbf{W}\mathbf{y}(0). \tag{34b}$$

As a key finding of the current paper, the matrix $\mathbf{C}(t)\mathrm{e}^{\mathbf{A}t}\mathbf{W} \in \mathbb{R}^{n\times n}$ is an approximation of the fundamental solution matrix $\mathbf{\Phi}(t)$, which is the matrix that satisfies $\mathbf{y}(t) = \mathbf{\Phi}(t)\mathbf{y}(0)$. In particular, for $t = T$, the monodromy matrix is approximated via

$$\mathbf{\Phi}_T \approx \mathbf{C}(T)\mathrm{e}^{\mathbf{A}T}\mathbf{W}. \tag{35}$$

If the Hill matrix $\mathbf{H}$ is already known due to other computations, e.g., as a by-product of a frequency-based continuation method such as MANLAB [27], then the similarity transform can be substituted into the matrix exponential to yield

$$\mathbf{\Phi}_T \approx \mathbf{C}(T)\mathbf{U}^\mathrm{T}\mathrm{e}^{\mathbf{H}T}\mathbf{U}\mathbf{W} =: \tilde{\mathbf{C}}(T)\mathrm{e}^{\mathbf{H}T}\tilde{\mathbf{W}} , \tag{36}$$

where $\tilde{\mathbf{C}}$, $\tilde{\mathbf{W}}$ can also be computed directly via

$$\tilde{\mathbf{W}} = \begin{pmatrix} \mathbf{I}_{n\times n} \\ \vdots \\ \mathbf{I}_{n\times n} \end{pmatrix} \tag{37a}$$

$$\tilde{\mathbf{C}} = \begin{pmatrix} \mathbf{0} \dots \mathbf{0} \; \mathbf{I}_{n\times n} \; \mathbf{0} \dots \mathbf{0} \end{pmatrix} \tag{37b}$$

by making use of the permutation properties of $\mathbf{U}$ (see Thm. 3). The naive choice for $\tilde{\mathbf{C}}$ was employed here for demonstration purposes, although all other choices are also applicable. Equations (35)–(37b) show that the choice of basis function ordering, and thus the exact numerical contents of the matrix, differ only formally and can be easily transformed into each other. For the sake of clarity, only the approximation (35) will be used in the following sections, unless otherwise stated. All results can, however, be transferred analogously to the formulation (4.1).

4.2 Stability and computational effort

As (35) is an approximation of the monodromy matrix, the Floquet multipliers can be approximated by its eigenvalues. This constitutes a novel projection-based stability method based on the Hill matrix. It is labeled *sorting-free* because the approach does not use the same steps as standard stability methods based on the Hill matrix [28,30,31].

In all standard approaches, the complete set of eigenvalues of the Hill matrix are determined in the first step. This is a computationally intense operation which may lack accuracy [41]. The result of this first operation is $N = n(2N_\mathbf{u} + 1)$ Floquet exponent candidates, of which only n are the best approximations to the true Floquet exponents. There are various sorting algorithms that vary in computational expense, which aim to find these n best approximations as introduced in Sect. 2.4.

In contrast to these approaches, we now propose a sorting-free method. The Hill matrix is constructed identically to the aforementioned approaches. However, once the Hill matrix is obtained, the sorting-free projection-based stability method goes a different route. The approximation to the monodromy matrix is determined first via (35). In the most general case, a matrix exponential is of the same complexity range $\mathcal{O}(N^3)$ as an eigenvalue problem [41,45], however this is not the case for the specific operations presented here. We give two reasons:

1. The matrix exponential in (35) is always multiplied by the matrix $\mathbf{W}$, which is smaller and relatively sparse. The matrix exponential can thus be computed by multiple evaluations of the action of the matrix exponential on a sparse vector. For this operation, efficient scaling and squaring approaches which utilize these properties exist [46].
2. In many applications the constructed Hill matrix will be relatively sparse. This is the case if the fre-

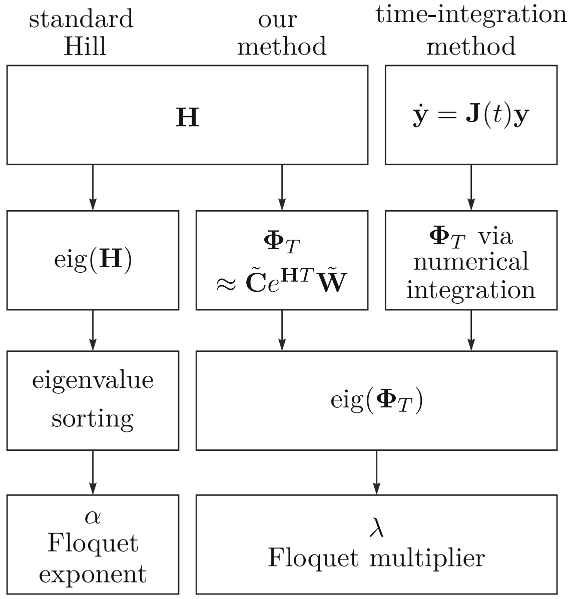

standard Hill
our method
time-integration method
$\mathbf{H}$
$\dot{\mathbf{y}} = \mathbf{J}(t)\mathbf{y}$
eig($\mathbf{H}$)
$\boldsymbol{\Phi}_T \approx \tilde{\mathbf{C}} e^{\mathbf{H}T} \tilde{\mathbf{W}}$
$\boldsymbol{\Phi}_T$ via numerical integration
eigenvalue sorting
eig($\boldsymbol{\Phi}_T$)
α Floquet exponent
λ Floquet multiplier

Fig. 2 Flowchart comparing the three general stability approaches. For each of the methods, the most computationally intense step is located in the second row

quencies of the dominant harmonics of the original system are small in comparison to the frequency order. This sparsity in $\mathbf{H}$, or equivalently in $\mathbf{A}$, can be exploited in the scaling and squaring algorithm, in addition to sparsity in $\mathbf{W}$.

As a second step in the sorting-free approach, it remains to solve an eigenvalue problem for the approximation of $\boldsymbol{\Phi}_T$. This matrix is only of the size $n \times n$, i.e., in general much smaller than the size $n(2N_{\mathbf{u}}+1) \times n(2N_{\mathbf{u}}+1)$ of the Hill matrix $\mathbf{H}$. The result are n Floquet multipliers and no a posteriori sorting of candidates is necessary. The reduction of candidates is performed implicitly by the projection matrices $\mathbf{C}$ and $\mathbf{W}$. This reduction through projection is essentially different from sorting, as the resulting Floquet multipliers from the projection-based method do generally not coincide identically with any Floquet multiplier candidates obtained as eigenvalues of the Hill matrix, transformed from Floquet exponents to Floquet multipliers via (15).

Alternatively to the Hill matrix approaches, the monodromy matrix can be determined via time-integration of the variational equation (14) [29]. Afterward, it again remains to solve the eigenvalue problem on the $n \times n$ monodromy matrix. The three general approaches are visualized in Fig. 2. While the starting point for the novel method is the Hill matrix as in the standard Hill methods, the final steps of our method are instead identical to the time-integration method. However, the second indicated step, which needs the most computational effort in all three approaches, is different. The presented sorting-free projection method therefore has the advantage of being a Hill-based method, which is favorable in an HBM setting, and at the same time only requiring to compute the smaller eigenvalue problem of the monodromy matrix, similar to the time-integration-based method.

4.3 Choice of projection matrix

Until now, it has been established that $\mathbf{C}(T)\mathrm{e}^{\mathbf{A}T}\mathbf{W}$ approximates the monodromy matrix if $\mathbf{C}\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}} = \mathbf{y}$. However, the choice of the matrix $\mathbf{C}$ to achieve this is not unique. In addition to the naive choice (31), many more matrices are admissible. Recall that the vector $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}$ (26) is sorted such that it contains n consecutive blocks of length $2N_{\mathbf{u}}+1$, each block collecting all functions of a single state. Therefore, the l-th row of $\mathbf{C}$, which singles out the l-th state, consists of $n-1$ blocks of zeros, each of length $2N_{\mathbf{u}}+1$, and one block $\mathbf{c}_l(t) \in \mathbb{R}^{1\times(2N_{\mathbf{u}}+1)}$ that is possibly filled with nonzero entries, yielding a non-square block-diagonal structure

$$\mathbf{C}(t) = \begin{pmatrix} \mathbf{c}_1(t) & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{c}_2(t) & \dots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \dots & \mathbf{c}_n(t) \end{pmatrix} \in \mathbb{C}^{n\times n(2N_{\mathbf{u}}+1)} \tag{38}$$

with $\mathbf{c}_l(t) \in \mathbb{C}^{1\times(2N_{\mathbf{u}}+1)}$; $l = 1, \dots, n$. With this notation, the requirement $\mathbf{C}(t)\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{y}, t) = \mathbf{y}$ for all $t \in [0, T]$ can be simplified to

$$\mathbf{c}_l(t)\mathbf{u}(t) = 1 \quad \forall t \in [0, T];\ l \in 1, \dots, n \tag{39}$$

by noting that y_l can be pulled out on both sides of the equation. As all entries in $\mathbf{u}$ are linearly independent, the time dependency of $\mathbf{c}_l$ is uniquely given via

$$\mathbf{c}_l(t) = \hat{\mathbf{c}}_l \begin{pmatrix} \mathrm{e}^{iN_{\mathbf{u}}\omega t} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \mathrm{e}^{-iN_{\mathbf{u}}\omega t} \end{pmatrix} =: \hat{\mathbf{c}}_l \mathbf{V}(t) \tag{40}$$

to cancel out the time dependency in (39), such that $\mathbf{V}(t)\mathbf{u}(t)$ is constant. In this expression, $\hat{\mathbf{c}}_l \in \mathbb{C}^{1\times(2N_{\mathbf{u}}+1)}$

is a constant row vector and $\mathbf{V}(t)$ is given by $\mathbf{V}(t) = \operatorname{diag}\left(e^{iN_{\mathbf{u}}\omega t}, \ldots, \; e^{-iN_{\mathbf{u}}\omega t}\right)$. Due to (39) and $\mathbf{y}$ being real, two additional conditions on $\hat{\mathbf{c}}_l$ are

$$\sum_{k=-N_{\mathbf{u}}}^{N_{\mathbf{u}}} \hat{c}_{l,k} = 1 \tag{41a}$$

$$\hat{c}_{l,k} = \bar{\hat{c}}_{l,-k} \; . \tag{41b}$$

All choices for $\hat{\mathbf{c}}_l$ that satisfy these conditions are admissible for the projection matrix. If a set $\left\{\hat{c}_{l,k}\right\}_{k=1}^{N_{\mathbf{u}}}$ of $N_{\mathbf{u}}$ arbitrary independent complex positive-frequency coefficients is given, they can be easily extended to an admissible choice $\hat{\mathbf{c}}_l$ using the conditions (41), which implies that the admissible projection matrices form a $nN_{\mathbf{u}}$-parameter family. For the sake of readability, the constraints will be left explicit in the considerations below. In Fig. 6–8 of Sect. 5.2, two choices for $\hat{\mathbf{c}}$ are compared and it turns out that this choice indeed strongly influences the approximation quality.

For numerical computations, handling the sparse matrix $\mathbf{C}$ with only few unknown coefficients is cumbersome. It is easier to collect all unknown variables $\hat{\mathbf{c}}_l, l = 1, \ldots, n$ into a long row vector $\hat{\mathbf{c}}_{\text{all}}$ with

$$\hat{\mathbf{c}}_{\text{all}} = \left(\hat{\mathbf{c}}_1 \; \ldots \; \hat{\mathbf{c}}_n\right) . \tag{42}$$

If the true monodromy matrix $\mathbf{\Phi}_T$ would be known, then (35) can be viewed as a fitting problem for the unknown $\hat{\mathbf{c}}_{\text{all}}$. The squared matrix residual

$$L_{\text{true}}(\hat{\mathbf{c}}_{\text{all}}) = \left\| \mathbf{C}(T) e^{\mathbf{A}T} \mathbf{W} - \mathbf{\Phi}_T \right\|^2 \tag{43}$$

should be minimized. Setting $\mathbf{Q} = e^{\mathbf{A}T}\mathbf{W}$ and utilizing the block diagonal structure (38) of $\mathbf{C}(T)$, this expression decouples row-wise into n (nonlinearly) constrained least-squares problems

$$\min_{\mathbf{c}_l(T)} \quad \left\| \mathbf{c}_l(T)^{\mathrm{T}} \left(\mathbf{Q}\right)_{l\text{-th block}} - \left(\mathbf{\Phi}_T\right)_{l\text{-th row}} \right\|^2 \tag{44a}$$

$$\text{subject to} \quad \sum_{k=-N_{\mathbf{u}}}^{N_{\mathbf{u}}} \hat{c}_{l,k} = 1 \tag{44b}$$

$$\hat{c}_{l,k} = \bar{\hat{c}}_{l,-k} \tag{44c}$$

for $l = 1, \ldots, n$. In this equation, the matrix $\mathbf{Q} \in \mathbb{C}^{n(2N_{\mathbf{u}}+1)\times(2N_{\mathbf{u}}+1)}$ is divided into n square blocks and due to the diagonal structure of $\mathbf{C}(T)$, only the l-th block of $\mathbf{Q}$ influences the l-th row of the least-squares problem (43). A matrix $\mathbf{C}$ that was determined for comparison purposes using this least-squares problem under knowledge of the true monodromy matrix will be referred to as $\mathbf{C}_{\text{true}}$ below. This least-squares problem has n equations for $N_{\mathbf{u}}$ independent unknowns. Therefore, to enable an optimal solution, it is advisable that the frequency order $N_{\mathbf{u}}$ is chosen to be at least n.

As the aim of this method is approximating the true monodromy matrix, which is unknown in applications, the least-squares problem (44) cannot be utilized in practice. As a practicable condition, the matrix $\mathbf{C}(t)$ can be chosen such that it optimally satisfies the variational equation (14). The residual of (14) for the approximated fundamental matrix is

$$\mathbf{R}(t, \hat{\mathbf{c}}_{\text{all}}) = \left(\mathbf{C}(t) e^{\mathbf{A}t} \mathbf{W}\right)^{\cdot} - \mathbf{J}(t)\mathbf{C}(t)e^{\mathbf{A}t}\mathbf{W} \tag{45}$$

and a cost function can be defined via

$$L_{\text{var}}(\hat{\mathbf{c}}_{\text{all}}) = \int_0^T \left\| \mathbf{R}(t, \hat{\mathbf{c}}_{\text{all}}) \right\|^2 \mathrm{d}t. \tag{46}$$

Similarly to the block-based approach for the least-squares problem, this function with a matrix argument can be transformed into a quadratic cost function

$$L_{\text{var}}(\hat{\mathbf{c}}_{\text{all}}) = \hat{\mathbf{c}}_{\text{all}} \left(\int_0^T \mathbf{\Theta}(t) \mathrm{d}t \right) \hat{\mathbf{c}}_{\text{all}}^{*} \tag{47}$$

with the vector $\hat{\mathbf{c}}_{\text{all}}$ as argument. This transformation of $\|\mathbf{R}(t)\|$ is detailed below.

By product rule, the total time derivative of the first summand in (45) yields

$$\left(\mathbf{C}(t)\mathbf{Q}(t)\right)^{\cdot} = \mathbf{C}(t)\left(\mathbf{A} + \mathbf{D}(t)\right)\mathbf{Q}(t) =: \mathbf{C}(t)\mathbf{L}(t) \; , \tag{48}$$

where $\mathbf{D} = \mathbf{I}_{n\times n} \otimes \operatorname{diag}(iN_{\mathbf{u}}\omega t, \ldots, -iN_{\mathbf{u}}\omega t)$ (cf. (40)) is the matrix that satisfies $\dot{\mathbf{C}} = \mathbf{CD}$ and $\dot{\mathbf{Q}} = \mathbf{AQ}$ follows from its definition. Below, the dependency on t in the matrices is omitted for the sake of brevity. Substitution of $\mathbf{Q}$, $\mathbf{L}$ into (45) yields

$$\mathbf{R} = \mathbf{CL} - \mathbf{JCQ}. \tag{49}$$

To exploit the diagonal structure (38) of $\mathbf{C}$, the matrices $\mathbf{Q}, \mathbf{L} \in \mathbb{C}^{n(2N_\mathbf{u}+1)\times(2N_\mathbf{u}+1)}$ are segmented into stacks of column vectors of length $2N_\mathbf{u} + 1$ via

$$\mathbf{L} =: \begin{pmatrix} \mathbf{L}_{11} & \dots & \mathbf{L}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{L}_{n1} & \dots & \mathbf{L}_{nn} \end{pmatrix} , \tag{50}$$

and $\mathbf{Q}$ analogously. For the first summand in (49), this gives

$$\mathbf{CL} = \begin{pmatrix} \mathbf{c}_1\mathbf{L}_{11} & \dots & \mathbf{c}_1\mathbf{L}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{c}_n\mathbf{L}_{n1} & \dots & \mathbf{c}_n\mathbf{L}_{nn} \end{pmatrix} , \tag{51}$$

where each of the indicated entries is a time-dependent scalar in $\mathbb{C}$. Similarly, the second summand can be separated into its entries to yield

$$\mathbf{JCQ} = \begin{pmatrix} \sum_{l=1}^{n} \mathbf{c}_l J_{1l}\mathbf{Q}_{l1} & \dots & \sum_{l=1}^{n} \mathbf{c}_l J_{1l}\mathbf{Q}_{ln} \\ \vdots & \ddots & \vdots \\ \sum_{l=1}^{n} \mathbf{c}_l J_{nl}\mathbf{Q}_{l1} & \dots & \sum_{l=1}^{n} \mathbf{c}_l J_{nl}\mathbf{Q}_{ln} \end{pmatrix} . \tag{52}$$

The still time-dependent $\mathbf{c}_l(t)$ is generated from the constant coefficients $\hat{\mathbf{c}}_l$ via the diagonal matrix $\mathbf{V}(t)$ as in (40). Collecting all unknown coefficients into one large vector $\hat{\mathbf{c}}_\text{all}$, the (i, j)-th scalar entry of the matrices in (49) can be expressed by

$$(\mathbf{CL})_{ij} = \hat{\mathbf{c}}_\text{all} \begin{pmatrix} \mathbf{0} & \dots & \mathbf{L}_{ij}^\mathrm{T}\mathbf{V} & \dots & \mathbf{0} \end{pmatrix}^\mathrm{T} =: \hat{\mathbf{c}}_\text{all}\mathbf{l}_{ij} \tag{53a}$$

$$(\mathbf{JCQ})_{ij} = \hat{\mathbf{c}}_\text{all} \begin{pmatrix} J_{i1}\mathbf{Q}_{1j}^\mathrm{T}\mathbf{V}, & \dots, & J_{in}\mathbf{Q}_{nj}^\mathrm{T}\mathbf{V} \end{pmatrix}^\mathrm{T} =: \hat{\mathbf{c}}_\text{all}\mathbf{q}_{ij} . \tag{53b}$$

In (53a), only the i-th block is nonzero.

If the Frobenius norm is used for the residual, all absolute squared values of the entries of the matrix (49) are summed, and this gives the expression

$$\|\mathbf{R}\|^2 = \sum_{i,j=1}^{n} \hat{\mathbf{c}}_\text{all} \left(\mathbf{l}_{ij} - \mathbf{q}_{ij}\right) \left(\mathbf{l}_{ij} - \mathbf{q}_{ij}\right)^* \hat{\mathbf{c}}_\text{all}^* . \tag{54}$$

Finally, noting that $\hat{\mathbf{c}}_\text{all}$ is not time-dependent, it can be pulled outside the integral to yield the quadratic cost function (47) with

$$\mathbf{\Theta}(t) = \sum_{i,j=1}^{n} \left(\mathbf{l}_{ij} - \mathbf{q}_{ij}\right) \left(\mathbf{l}_{ij} - \mathbf{q}_{ij}\right)^* . \tag{55}$$

Therefore, the search for the best choice for the projection matrix $\mathbf{C}_\text{var}$ can be formulated as a quadratic program

$$\min_{\hat{\mathbf{c}}_\text{all}} \quad \hat{\mathbf{c}}_\text{all} \left(\int_0^T \mathbf{\Theta}(t)\mathrm{d}t\right) \hat{\mathbf{c}}_\text{all}^* \tag{56a}$$

$$\text{subject to} \quad \hat{\mathbf{c}}_\text{all}\mathbf{W} = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix} , \tag{56b}$$

where the equality constraint encodes the normalization condition (41a). The condition (41b) is not explicitly stated in the quadratic program. This is because minimizers of (56) always fulfill this condition. From an arbitrary candidate $\hat{\mathbf{c}}_\text{all}$, one can construct a symmetric candidate $\hat{\mathbf{c}}_\text{sym}$ which fulfills (41b) and for which $\mathbf{R}$ has the same real part and zero imaginary part, meaning that $\|\mathbf{R}\|^2$ can not be larger for the symmetric candidate than for the non-symmetric one. The proof of this is sketched in Appendix B.3.

As the matrix $\mathbf{\Theta}(t)$ is of size $n(2N_\mathbf{u}+1)\times n(2N_\mathbf{u}+1)$, and therefore rather large, efficient numerical determination of the integral (56a) is crucial to retain a numerically efficient stability method. Because the time-dependency of $\mathbf{\Theta}$ is introduced by terms of the form $\mathrm{e}^{ik\omega t}$, the integral can be reformulated into a (infinite) sum of inner products with Fourier base functions if the power series expression for the matrix exponential is used. This suggests that the integral could be computed efficiently using FFT methods. However, due to the matrix exponential there is no limit in frequency of these terms and aliasing effects must be considered.

Alternatively, the integral can be determined using numerical quadrature schemes. This is accurate, but computationally expensive. Finally, it is also an option to require the integrand to be zero only at specific time instants. This reduces the quadratic program to a linear equation system. While this approach does not minimize the quadratic program (56) in an integral sense, the resulting approximate monodromy matrix seems to be relatively accurate in application.

5 Numerical examples

The theoretical results of the previous sections will be illustrated in this section using some numerical examples, which will allow us to demonstrate the numerical efficiency and accuracy of the proposed projection-based Hill method. The Mathieu equation is utilized as a simple linear time-periodic system of the form (14) to explicitly illustrate the Koopman lift and the projection-based stability approach. To demonstrate the computational advantage for larger numerical orders, the linearization of the vertically excited n-pendulum is then considered as a generalization of the Mathieu equation with arbitrary degrees of freedom.

5.1 Mathieu equation

The Mathieu equation

$$\ddot{x} + (a + 2b\cos 2\omega t)x = 0 \tag{57}$$

is an example of a Hill differential equation [47], which has become very well-known since it results from linearization of a number of applications, among them rolling of container ships [48] and a vertically excited pendulum [49]. After bringing the system into first-order-form

$$\dot{\mathbf{x}}(t) = \mathbf{J}(t)\mathbf{x}(t) = \begin{pmatrix} 0 & 1 \\ -a - 2b\cos(2\omega t) & 0 \end{pmatrix}\mathbf{x}(t) \tag{58}$$

with $\mathbf{x} = [x, \dot{x}]^{\mathrm{T}}$, (58) is a linear periodic time-varying homogeneous system of the form (12) with system period $\tilde{T} = \frac{\pi}{\omega}$, meaning that it is suitable to explore the stability methods of Sect. 4. Necessarily, the system is also T-periodic with $T = 2\tilde{T} = \frac{2\pi}{\omega}$. It is known that the only parameter combinations of (58) which admit non-trivial T-periodic solutions are located at the stability boundaries [39,50]. The stability boundaries of the Mathieu equation can therefore be identified using the HBM or the shooting method. Below, the base frequency ω (i.e., including the first subharmonics of the system) has been chosen for investigation using the frequency-based methods. This is due to two reasons: First, as the stability boundaries may admit T-periodic or $\frac{T}{2}$-periodic solutions, this choice allows to identify all stability boundaries using one unified HBM approach without further distinction. Second, in the projection-based Hill method, this choice of base frequency allows to better showcase the influence of the projection matrix.

The equivalence of the Koopman lift matrices, the Hill matrix and the HBM are explicitly verified for the smallest possible frequency order $N_{\mathbf{u}} = 1$. First, the Hill matrix is constructed in the standard way. The Fourier decomposition of $\mathbf{J}(t)$ can be read directly from (58). The nonzero Fourier coefficients with base frequency ω are

$$\mathbf{J}_0 = \begin{pmatrix} 0 & 1 \\ -a & 0 \end{pmatrix} \tag{59}$$

$$\mathbf{J}_2 = \mathbf{J}_{-2} = \begin{pmatrix} 0 & 0 \\ -b & 0 \end{pmatrix} \tag{60}$$

and all others vanish. By construction using (18), the Hill matrix of frequency order 1 reads

$$\mathbf{H} = \begin{pmatrix} i\omega & 1 & 0 & 0 & 0 & 0 \\ -a & i\omega & 0 & 0 & -b & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -i\omega & 1 \\ -b & 0 & 0 & 0 & -a & -i\omega \end{pmatrix}. \tag{61}$$

The block-Toeplitz-like structure with additional diagonal elements is visible in (61). Furthermore, it is obvious through the parameter b that terms of harmonic 2 influence the Hill matrix, even if the frequency order was chosen to be $N_{\mathbf{u}} = 1 < 2$.

Further, assume there exists a (potentially non-trivial) periodic solution $\mathbf{x}_p = \sum_j \mathbf{p}_j \mathrm{e}^{ij\omega t}$. Because the original dynamics (58) is already linear, this also holds for the perturbed dynamics

$$\dot{\mathbf{y}}(t) = \mathbf{J}(t)\mathbf{y}(t) + \mathbf{J}(t)\mathbf{x}_p(t) - \dot{\mathbf{x}}_p(t). \tag{62}$$

With (62) expressed explicitly in Fourier series form

$$\dot{y}_1 = y_2 + \sum_j (p_{j,2} - ij\omega p_{j,1})\mathrm{e}^{ij\omega t} \tag{63a}$$

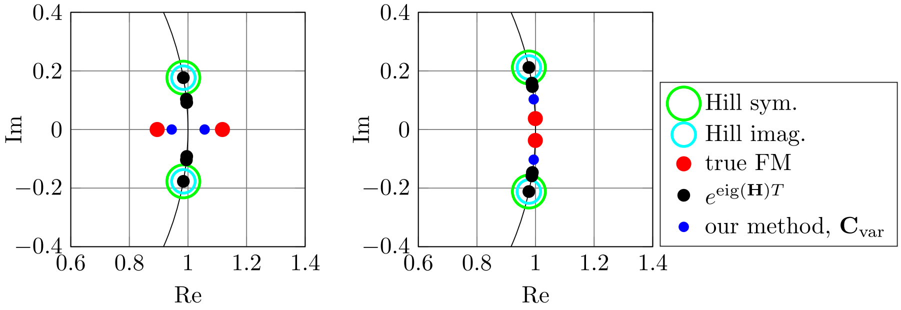
$$\begin{aligned}\dot{y}_2 = &-ay_1 - by_1\mathrm{e}^{-2i\omega t} - by_1\mathrm{e}^{2i\omega t} \\ &+ \sum_j \left[-ap_{j,1} - b(p_{j+2,1} + p_{j-2,1}) - ij\omega p_{j,2}\right]\mathrm{e}^{ij\omega t},\end{aligned} \tag{63b}$$

the explicit determination of the time derivatives of the Koopman basis functions (26) for the perturbed dynamics yields with the index shift $\tilde{j} = j + k$

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}[y_1\mathrm{e}^{ik\omega t}] &= \dot{y}_1\mathrm{e}^{ik\omega t} + ik\omega y_1\mathrm{e}^{ik\omega t} \\ &= y_2\mathrm{e}^{ik\omega t} + ik\omega y_1\mathrm{e}^{ik\omega t} \\ &\quad + \sum_{\tilde{j}}\left[p_{(\tilde{j}-k,2)} - i(\tilde{j}-k)\omega p_{(\tilde{j}-k,1)}\right]\mathrm{e}^{i\tilde{j}\omega t}\end{aligned} \tag{64a}$$

for the first state and

Fig. 3 Floquet multiplier locations for $\omega = 1$ and $b = 1.2$, determined using a Hill matrix of order 4



(a) $a = -0.367$. (b) $a = -0.3673$.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}[y_2 \mathrm{e}^{ik\omega t}] = &- a y_1 \mathrm{e}^{ik\omega t} + ik\omega y_2 \mathrm{e}^{ik\omega t} \\ &- b y_1 \mathrm{e}^{i(k-2)\omega t} - b y_1 \mathrm{e}^{i(k+2)\omega t} \\ &+ \sum_{\tilde{j}} \left[-a p_{(\tilde{j}-k),1} - i\omega(\tilde{j}-k) p_{(\tilde{j}-k),2} \right] \mathrm{e}^{i\tilde{j}\omega t} \\ &- b \sum_{\tilde{j}} \left[p_{(\tilde{j}+2-k),1} + p_{(\tilde{j}-2-k),1} \right] \mathrm{e}^{i\tilde{j}\omega t}\end{aligned} \tag{64b}$$

for the second state. The coefficients for the basis functions can be identified in (64) by inspection. For the state-dependent basis functions with $N_{\mathbf{u}} = 1$, this yields

$$\mathbf{A} = \begin{pmatrix} -i\omega & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & i\omega & 0 & 0 & 1 \\ & & & & & \\ -a & 0 & -b & -i\omega & 0 & 0 \\ 0 & -a & 0 & 0 & 0 & 0 \\ -b & 0 & -a & 0 & 0 & i\omega \end{pmatrix}. \tag{65}$$

The blocks of $\mathbf{A}$ have been visually separated to indicate the dependence on the individual states. The selection matrix

$$\mathbf{U} = \begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix} \tag{66}$$

necessary to reorder the basis $\mathbf{\Psi}_{\mathbf{z},\mathrm{lin}}$ as required by Theorem 3 can be determined by inspection. With this, it is indeed easy to verify that $\mathbf{H} = \mathbf{U}\mathbf{A}\mathbf{U}^{\mathrm{T}}$ holds for Eqs. (61), (65), (66).

The state-independent terms in (64) can be collected into the $\mathbf{B}$ matrix. For the considered case $N_{\mathbf{u}} = 1$, only the values $k, j \in \{-1, 0, 1\}$ are considered. As expected from Theorem 2, this yields two stacked Toeplitz matrices $\mathbf{B}^{\mathrm{T}} = [\mathbf{B}_1^{\mathrm{T}}, \mathbf{B}_2^{\mathrm{T}}]$ with

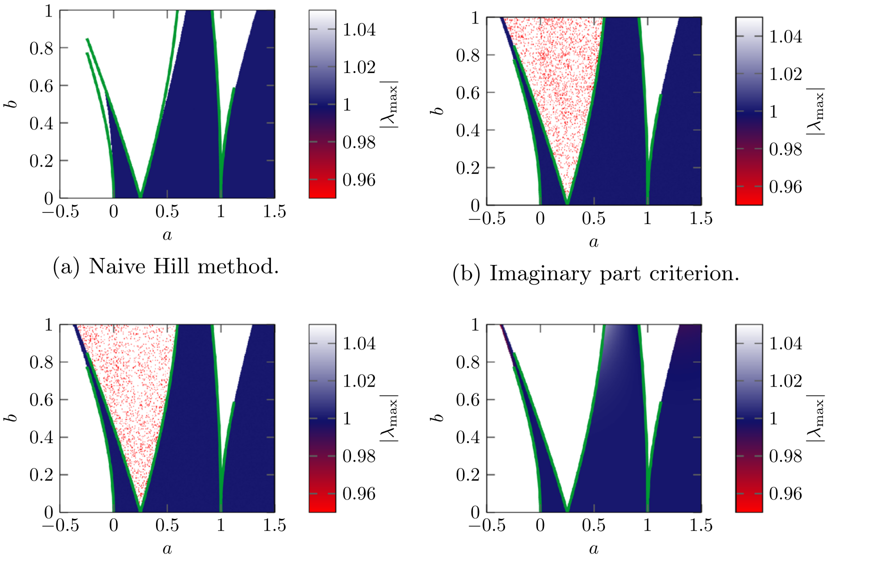
$$\mathbf{B}_1 = \begin{pmatrix} p_{0,2} & p_{1,2} - i\omega p_{1,1} & p_{2,2} - 2i\omega p_{2,1} \\ p_{-1,2} + i\omega p_{-1,1} & p_{0,2} & p_{1,2} - i\omega p_{1,1} \\ p_{-2,2} + 2i\omega p_{-2,1} & p_{-1,2} + i\omega p_{-1,1} & p_{0,2} \end{pmatrix} \tag{67}$$

and $\mathbf{B}_2$ omitted for the sake of brevity. The expected Toeplitz structure and the HBM equations up to order $2N_{\mathbf{u}}$ are clearly visible. It is also easy to see that $\mathbf{B}_1$ holds the HBM equations up to order 2 for the first (time-independent) row of (58).

The accuracy of a stability traverse is investigated in Fig. 3 for the Mathieu equation with $\omega = 1$ and $b = 1.21$. The parameter a takes the values -0.367 and -0.3673, respectively. Between these values, a stability change from unstable to stable occurs, where the pair of Floquet multipliers meet at 1, and continue on the unit circle as a complex conjugated pair. The Floquet multipliers labeled as true in Fig. 3 were determined using the time-integration method with a high relative and absolute tolerance of 10^{-14} each. In contrast, all Hill-matrix-based methods rely on a Hill matrix of relatively low frequency order $N_{\mathbf{u}} = 4$, such that the differences in performance are more visible.

From the figures, it is apparent that the projection-based method with an optimized projection matrix $\mathbf{C}_{\mathrm{var}}$ (i.e., by solving (56)) can be able to track the true Floquet multipliers more closely than any Floquet multiplier candidates obtained directly from the eigenval-

Fig. 4 Ince–Strutt diagrams created with various methods based on the Hill matrix of order $N_{\mathbf{u}} = 4$. Color indicates absolute value of largest considered Floquet multiplier candidate. Green lines indicate true stability boundaries



(a) Naive Hill method.

(b) Imaginary part criterion.

(c) Symmetry criterion.

(d) Novel projection-based method with naive **C** (31).

ues of the Hill matrix allow. Even more, both considered sorting methods choose another pair of eigenvalues than the closest one.

The stability regions of the Mathieu equation are often visualized in a so-called Ince–Strutt diagram [51]. Figure 4 showcases the properties of the different approaches for drawing this stability map using the Hill matrix of a relatively low frequency order $N_{\mathbf{u}} = 4$. The color indicates the absolute value of the largest Floquet multiplier (in magnitude). Dark blue colors indicate a maximum value of exactly 1, i.e., stable regions. White color indicates a maximum value larger than 1, i.e., unstable regions.

For Hill equations, it is known that the product of both Floquet multipliers must always equal to 1 [39], i.e., the largest Floquet multiplier may never admit an absolute value smaller than one. In the diagrams, red and purple regions correspond to parameter combinations where the largest identified Floquet multiplier has an absolute value smaller than one, meaning that the identified Floquet multipliers can certainly not be the true ones. In every figure, the true stability boundaries are given in green by the solution of an accurate shooting method.

If all eigenvalues of the Hill matrix are considered for stability (naive approach, Fig. 4a), there are large regions that are wrongly classified as unstable, while none of the unstable regions are wrongly classified as stable. This is expected since in the unstable case, the best approximations will always be among the considered eigenvalues, but in the stable case the additional candidates add additional possibilities of asserted instability. In contrast, the classical Hill method with sorting procedures, which is the current state-of-the-art, classifies most of the stable regions correctly. This is visualized in Fig. 4b for the imaginary-part-based criterion [23,30] and in Fig. 4c for the symmetry-based criterion [27,28]. However, in the instability tongue that is separated by non-trivial solutions of period T, red artifacts that are classified as stable are visible in both approaches. These correspond to cases where the sorting algorithm did not choose the correct Floquet multipliers (of which one would be stable, i.e., real-valued and < 1, and the other unstable, i.e., real-valued and > 1), but rather chose two instances of the stable class, both with real parts < 1. With the naive projection matrix, the novel projection-based approach preserves most of the stability regions, while not exhibiting any stable artifacts. However, at the edges of the stable regions, in particular around $a = 0.5$ and $b = 1$, slightly larger values of the magnitude are visible. The stable regions are slightly underestimated. A very similar behavior can also be observed with the optimized projection matrix $\mathbf{C}_{\text{var}}$. The example with the naive projection matrix shows that the presented method can be very accurate, even in its simplest form which is easy to implement since only the matrix equation (4.1) is needed.

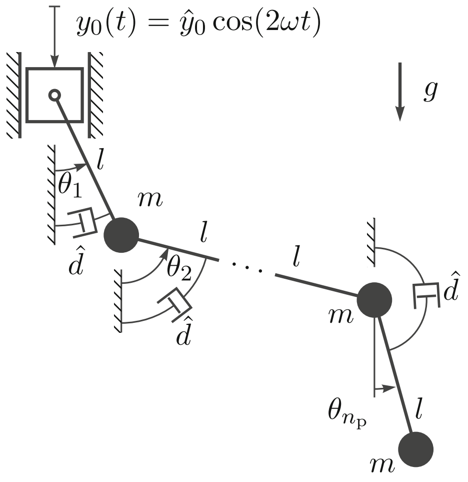


Fig. 5 The vertically excited multiple pendulum

5.2 Vertically excited multiple pendulum

The Mathieu equation considered in Sect. 5.1 can result from linearization of a vertically excited mathematical pendulum, also called the Kapitza pendulum [52]. As a scalable generalization for arbitrary degrees of freedom, the linearized dynamics of a vertically excited multiple pendulum is considered. A sketch of the considered mechanical system is given in Fig. 5. The pendulum consists of n_p joints, each of mass m, with viscous absolute damping $\hat{d}$, linked by n_p rods of length l. The minimal coordinates $\boldsymbol{\theta} = \left(\theta_1, \ldots, \theta_{n_\mathrm{p}}\right)^\mathrm{T}$ are the absolute angles of the individual joints. The suspension point of the pendulum moves vertically with $y_0(t) = \hat{y}_0 \cos(2\omega t)$. Gravitation acts in the vertical direction.

The equations of motion for the vertically excited pendulum can be derived similar to [53]. We define the auxiliary vectors

$$\mathbf{s}(\boldsymbol{\theta})^\mathrm{T} := \left(n_\mathrm{p} \sin\theta_1, (n_\mathrm{p}-1)\sin\theta_2, \ldots, \sin\theta_{n_\mathrm{p}}\right) \tag{68a}$$

$$\mathbf{c}(\boldsymbol{\theta})^\mathrm{T} := \left(n_\mathrm{p} \cos\theta_1, (n_\mathrm{p}-1)\cos\theta_2, \ldots, \cos\theta_{n_\mathrm{p}}\right) \tag{68b}$$

$$\dot{\boldsymbol{\theta}}^2 := \left(\dot{\theta}_1^2, \ldots, \dot{\theta}_{n_\mathrm{p}}^2\right)^\mathrm{T} \tag{68c}$$

as well as matrices $\mathcal{S}(\boldsymbol{\theta}), \mathcal{C}(\boldsymbol{\theta}) \in \mathbb{R}^{n_\mathrm{p} \times n_\mathrm{p}}$ with

$$\mathcal{S}_{ij}(\boldsymbol{\theta}) := \left[n_\mathrm{p} + 1 - \max(i, j)\right] \sin(\theta_i - \theta_j) \tag{68d}$$

$$\mathcal{C}_{ij}(\boldsymbol{\theta}) := \left[n_\mathrm{p} + 1 - \max(i, j)\right] \cos(\theta_i - \theta_j). \tag{68e}$$

Dropping the arguments for the sake of brevity, the potential energy $V(\boldsymbol{\theta}, t)$ and the kinetic energy $T(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, t)$ are given by

$$V = -n_\mathrm{p} m g y_0 - mgl \left(1 \, \ldots \, 1\right) \mathbf{c} \tag{69a}$$

$$T = \frac{1}{2} m \left[n_\mathrm{p} \dot{y}_0^2 + l^2 \dot{\boldsymbol{\theta}}^\mathrm{T} \mathcal{C} \dot{\boldsymbol{\theta}} - 2 l \dot{y}_0 \mathbf{s}^\mathrm{T} \dot{\boldsymbol{\theta}}\right]. \tag{69b}$$

Using the Lagrange equations of the second kind (see, e.g., [54, p. 76]), the equations of motion are given by

$$\mathcal{C}\ddot{\boldsymbol{\theta}} + \frac{\hat{d}}{ml^2}\dot{\boldsymbol{\theta}} + \mathcal{S}\dot{\boldsymbol{\theta}}^2 + \left(-\frac{\ddot{y}_0(t)}{l} + \frac{g}{l}\right)\mathbf{s}(\boldsymbol{\theta}) = \mathbf{0} \tag{70}$$

after some algebra. Introducing the abbreviations

$$\mathbf{M} := \mathcal{C}(\mathbf{0}) \tag{71a}$$

$$\mathbf{D} := \operatorname{diag}(n_\mathrm{p}, \ldots, 1) \tag{71b}$$

$$a := g/l \tag{71c}$$

$$2b := 4\hat{y}_0\omega^2/l \tag{71d}$$

$$d := \frac{\hat{d}}{ml^2} \tag{71e}$$

and linearizing around the origin, the first-order linearized dynamics of the vertically excited pendulum is

$$\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{M} \end{pmatrix} \begin{pmatrix} \dot{\boldsymbol{\theta}} \\ \ddot{\boldsymbol{\theta}} \end{pmatrix} = \begin{pmatrix} \mathbf{0} & \mathbf{I} \\ -[a + 2b\cos(2\omega t)]\mathbf{D} & -d\mathbf{I} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta} \\ \dot{\boldsymbol{\theta}} \end{pmatrix}, \tag{72}$$

which, after inversion of the left matrix, is of the form $\dot{\mathbf{y}} = \mathbf{J}(t)\mathbf{y}$ that can be analyzed using the presented methods. The total number of system states is $n = 2n_\mathrm{p}$. The relation between the classical Mathieu equation (58) and the linearized vertically excited multiple pendulum is visible from (72). The total number n of states as considered throughout this paper is $n = 2n_\mathrm{p}$.

To analyze the convergence behavior of our proposed method, the accuracy of the Floquet multipliers of the equilibrium for one specific parameter set (a, b, d) is evaluated for various truncation orders $N_\mathbf{u}$ and using the various approaches discussed in this paper. As a basis for comparison, the "true" Floquet multipliers are determined by integrating the variational equation (14) using the MATLAB function `ode45` with absolute and relative tolerances of

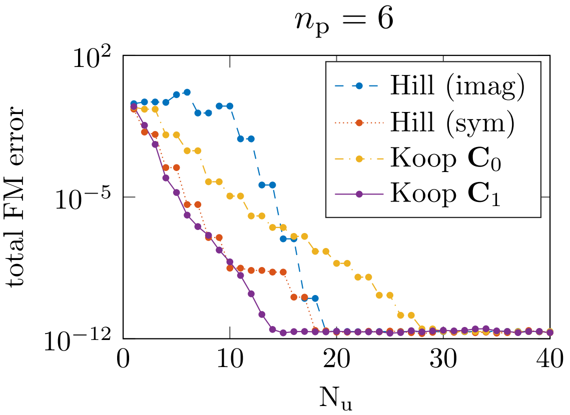


Fig. 6 Accuracy of Floquet multipliers against Floquet multipliers by time-integration over frequency order for the linearized 6-pendulum with $(a, b, d) = (5, 0.5, 0.2)$

10^{-13}. The total Floquet multiplier error is defined as the square root of the sum of the squared absolute differences between the true Floquet multipliers and the obtained Floquet multipliers, while the latter are ordered such that this error is minimal. More formally, let $\mathcal{P} := \{\pi : \{1, \ldots, n\} \to \{1, \ldots, n\} \,|\, \pi \text{ bijective}\}$ be the (finite) set of permutations, i.e., reorderings, of the index set $\{1, \ldots, n\}$. Then, the total Floquet multiplier error is given by

$$\varepsilon_\mathrm{total} = \min_{\pi \in \mathcal{P}} \sqrt{\sum_{l=1}^{n} \left| \lambda_{l,\mathrm{true}} - \lambda_{\pi(l),\mathrm{cand}} \right|^2}. \tag{73}$$

For the two standard Hill approaches, the eigenvalues and eigenvectors of the Hill matrix $\mathbf{H}$ were determined using the MATLAB procedure `eig`. The sorting procedure based on the imaginary part as described in [23,26,30] then singles out the n eigenvalues with least imaginary part in modulus, while the symmetry-based sorting procedure singles out the n eigenvalues whose eigenvectors have lowest weighted mean according to [27]. The Floquet multipliers are then determined from the Floquet exponents (i.e., the identified eigenvalues) using (15). The corresponding errors are depicted in Fig. 6 in dashed blue and dotted red for the imaginary part criterion and the symmetry criterion, respectively.

The presented novel Koopman-based method is evaluated for two choices of the projection matrix $\mathbf{C}$ (see Sect. 4.3). The Floquet multipliers are determined as the eigenvalues of the approximate monodromy matrix (35). The matrix exponential in (35) is evaluated directly with its action onto the matrix $\mathbf{W}$ using the MATLAB function `exmpv` [46,55]. The dash-dotted yellow line shows the accuracy of the Floquet multipliers obtained from the monodromy approximation using the naive projection choice $\mathbf{C}_0$ (31), while the solid purple line indicates the accuracy obtained by a more informed projection matrix

$$\mathbf{C}_1 := \mathbf{I}_{n\times n} \otimes \left(1, -1, 1, -1, \ldots, -1, 1\right) , \tag{74}$$

which was obtained by inspection after running the optimization (56) for a few test cases. From the figure, it is visible that all considered approaches eventually converge to the true Floquet multipliers. The final error of order 10^{-12} can be attributed to inaccuracies of the numerical integration. The imaginary-part-based criterion, while being the only one with a rigorous convergence proof for $n \to \infty$, performs highly inaccurate for smaller $N_\mathbf{u}$. Moreover, the choice of projection matrix significantly influences the performance of the novel projection-based approach. While the naive choice of projection matrix does converge toward the correct value, the optimization-based projection matrix exhibits a significantly better convergence rate which can compete with the symmetry-based criterion or outperforms it.

The data set of Fig. 6 is again plotted in Fig. 7, but against the computation time instead of the frequency order, giving a work-precision diagram. The strength of the projection-based methods for higher frequency orders is apparent in this figure. While convergence of the imaginary-part-based method with respect to the frequency order $N_\mathbf{u}$ is better than that of the new projection-based method for the naive projection choice $\mathbf{C}_0$, the accuracy of the projection-based method is higher than that of the imaginary-part-based method for any computation time. The reason for this is that the eigenproblem becomes more expensive than the matrix exponential for higher matrix dimensions, i.e., for higher $N_\mathbf{u}$. The same property can also be observed regarding the symmetry-based approach and the novel projection-based approach with optimized projection matrix $\mathbf{C}_1$. The projection-based approach is able to achieve maximum accuracy already at approx. 0.1s compared to approx. 0.25s for the symmetry-based approach. This is a ratio of $\frac{2}{5}$, which is considerably lower than the corresponding ratio $\frac{15}{18}$ of

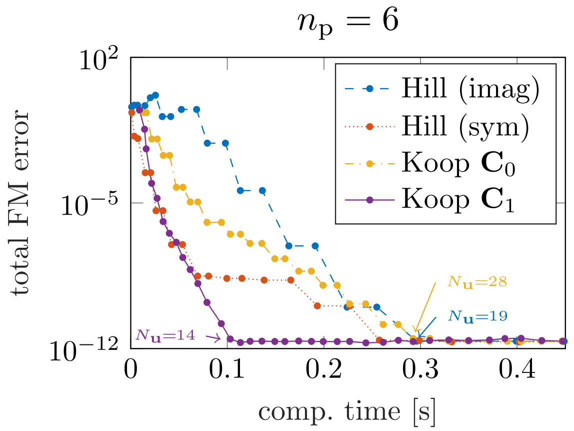


Fig. 7 Accuracy of Floquet multipliers over computation time for the linearized 6-pendulum with $(a, b, d) = (5, 0.5, 0.2)$. The data points correspond to values of $N_\mathbf{u}$ in the range between 1 and 40

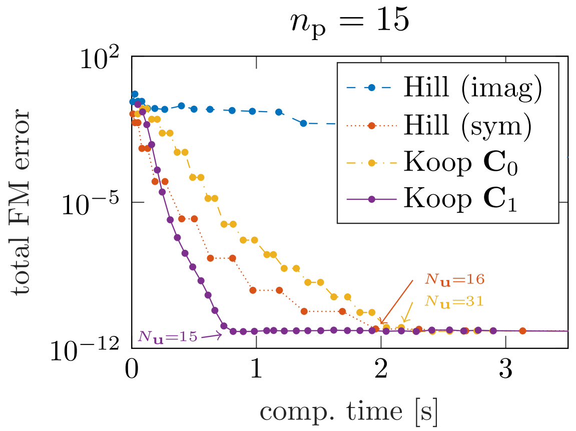


Fig. 8 Accuracy of Floquet multipliers over computation time for the linearized 15-pendulum with $(a, b, d) = (5, 0.5, 0.2)$. The data points correspond to values of $N_\mathbf{u}$ in the range between 1 and 40

frequency orders. If the system dimension increases, this efficiency difference increases as well as the size $n(2N_\mathbf{u} + 1) \times n(2N_\mathbf{u} + 1)$ of the Hill matrix is influenced by a product of n and $N_\mathbf{u}$. Therefore, the superior efficiency of the projection-based approaches against the standard approaches is even more prominent. Figure 8 again compares computation time against accuracy for the various approaches, but for a pendulum with 15 instead of 6 links. Since the number of states is increased, the computation time for all the considered approaches is larger than in Fig. 7. However, the increase in number of states impacts the standard Hill methods which rely on solving the large eigenvalue problem more than the projection-based method. This is because of the key feature of the new method: a moderate growth of computational cost as sparsity can be exploited in the action of the matrix exponential. In fact, in Fig. 8, the break-even point between the symmetry-based method and the projection-based method with the optimized projection matrix already occurs at $N_\mathbf{u} = 6$. The imaginary-part-based Hill method fails to converge within the depicted computation time interval.

6 Conclusions

In this paper, a relationship between the HBM, the Hill method and the Koopman lift with specific basis functions has been addressed. It has been shown that the Hill matrix is the system matrix of a linear time-invariant dynamical system of higher order which approximates the perturbed dynamics. This relationship has been used to derive a novel stability method based on projecting the Hill matrix down to original system size, instead of computing all its eigenvalues.

The resulting stability criterion in its naive form is remarkably simple to implement. It only needs a matrix exponential and two projection matrices, which are the same for all systems and easy to construct (see (31), (33)). It has been shown in the examples of Sect. 5 that this simple method can compete with the state-of-the-art algorithms in terms of Floquet multiplier accuracy over computational effort.

To further improve the accuracy of the proposed method, an optimization-based approach for the choice of the projection matrix through a quadratic program has been presented in Sect. 4.3. While the determination of the matrix integral in (56) is more computationally expensive, it has been shown in Sect. 5 that this optimized projection matrix allows for even better accuracy.

There are many topics for further research expanding from the results of this work: The approach itself offers opportunities for additional development both in its theoretical as well as for its computational properties. Further, there are many possible interesting applications of the approach to technical systems, and extensions to wider classes of systems would be valuable.

With respect to theoretical properties, there is currently no rigorous convergence proof for the novel approach for any choice of projection matrix. Such a convergence proof would aid in an *a priori* determination which method to employ for the best results.

Considering computational properties, while the presented approach does reduce computational effort compared to the state-of-the-art approaches, some bottlenecks remain. In particular, the efficient sparsity-promoting determination of the matrix exponential with projection from both sides as well as the integral in the optimization problem (56) would benefit from more efficient computation techniques. As the latter is constructed by products of the Fourier coefficients of $\mathbf{J}$ in $\mathbf{A}$ and additional Fourier terms, there could exist an operation to obtain this integral directly from the (FFT) Fourier series of $\mathbf{J}(t)$.

From an application point of view, going beyond the simple and mainly academical examples presented in this work, it would also be expedient to apply the novel stability method to systems of more practical relevance. To achieve this goal, it would be beneficial to integrate the novel method within an established continuation framework. The authors believe that the MANLAB continuation framework [25,28] would be especially suitable because the Hill matrix can be constructed in this framework without much additional effort. Large systems that were already analyzed extensively using the MANLAB framework (e.g., [56]) could benefit from the stability insights of our proposed method, while simultaneously serving as practical examples of the performance of the method.

Regarding possible extensions, while some relations were established in Theorem 1 for polynomial basis functions of the Koopman lift, most sections of this work rely on basis functions that are linear in the state. Some recent results [57,58] make statements about return time and period for autonomous systems based on the Carleman linearization. As our (full) basis provides a natural extension of the Carleman linearization, these results could possibly be integrated into the presented framework.

Finally, it would be of practical value to extend the method to a wider class of systems, for example, possibly including delay differential equations (DDEs) with delayed coordinates in the basis functions.

Author contributions Both authors contributed to the conceptualization, design and methodology. The formal analysis was performed by Fabia Bayer. The original draft of the manuscript was prepared by Fabia Bayer and both authors commented on previous versions of the manuscript. Remco Leine was responsible for funding acquisition and supervision. Both authors read and approved the final manuscript.

Funding Open Access funding enabled and organized by Projekt DEAL. The authors declare that no funds, grants, or other support was received during the preparation of this manuscript.

Data availability All numerical data in this work have been generated from the considered system Eqs. (58), (72), with the approaches described in this work and the cited works, using MATLAB standard methods and the MATLAB function `expmv` [55]. Therefore, it is possible to completely reproduce the data from the information given in this work, except for the computation times in the work-precision diagrams. These computation times have been evaluated on a Windows PC, and more detailed computation time data can be obtained from the authors upon request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Appendix A: Multi-index calculation rules

For convenience, the standard multi-index calculation rules (see, for example, [59, p. 319]) are revisited. A multi-index is a tuple $\boldsymbol{\beta} \in \mathbb{N}_0^d$ with nonnegative integer entries. In particular, a multi-index $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_d)$ has the 1-norm

$$\|\boldsymbol{\beta}\| = \sum_{i=1}^{d} \beta_i \tag{A1}$$

and a factorial

$$\boldsymbol{\beta}! = \prod_{i=1}^{d} \beta_i! \tag{A2}$$

given by the product of the factorials of all its components. For a vector $\mathbf{x} \in \mathbb{C}^d$, exponentiation is given by

$$\mathbf{x}^{\boldsymbol{\beta}} = \prod_{i=1}^{d} x_i^{\beta_i} \tag{A3}$$

and for a function $f : \mathbb{R}^d \to \mathbb{C}$, the $\boldsymbol{\beta}$-th derivative is given by

$$\frac{\partial^{\boldsymbol{\beta}} f}{\partial \mathbf{x}^{\boldsymbol{\beta}}} = \frac{\partial^{\|\boldsymbol{\beta}\|} f}{\partial x_1^{\beta_1} \dots \partial x_d^{\beta_d}}. \tag{A4}$$

It is known in combinatorics that the number of multi-indices $\boldsymbol{\beta} \in \mathbb{N}_0^d$ with $\|\boldsymbol{\beta}\| \leq N$ is given by $\binom{N+d}{d}$ [60]. This is also the number of unique monomials of the form (A3) with degree $\leq N$, including $\mathbf{x}^{\mathbf{0}}$.

Appendix B: Proofs for the theorems

In this section, the theorems of Sect. 3.2 are proven formally. The slightly unusual matrix-valued inner product notation (1)–(2c) is heavily utilized throughout the proofs for the sake of brevity.

B.1 Theorems 1 and 2

Theorem 2 will be proven first, and Theorem 1 will result as a consequence from that proof. Recall Theorem 2:

Theorem 2 *Let $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{p})\mathbf{z} + \mathbf{B}(\mathbf{p})\mathbf{u}$ be the lifted dynamics of system (6) with linear basis functions $\boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}$ of frequency order $N_{\mathbf{u}}$ that are sorted as in (22), evaluated for the perturbed system around an unknown periodic ansatz of the form (10) up to frequency order at least $N_{\mathrm{HBM}} = 2N_{\mathbf{u}}$. Then, the matrix $\mathbf{B}(\mathbf{p}) \in \mathbb{C}^{n(2N_{\mathbf{u}}+1)\times(2N_{\mathbf{u}}+1)}$ consists of n stacked Toeplitz matrices. The l-th Toeplitz matrix $\mathbf{B}_l$ contains as entries (ignoring duplicates) precisely the $4N_{\mathbf{u}} + 1$ residuals $r_{k,l}(\mathbf{p})$, $|k| \leq 2N_{\mathbf{u}}$ that follow from the HBM w.r.t the l-th state.*

If $\mathbf{B}(\mathbf{p}) = \mathbf{0}$, then all these residuals of the HBM vanish. Conversely, if $\mathbf{p}$ solves the HBM equations $\mathbf{r}_k(\mathbf{p}) = \mathbf{0}$, $|k| \leq 2N_{\mathbf{u}}$, then it holds that $\mathbf{B}(\mathbf{p}) = \mathbf{0}$.

Proof For the sake of brevity, bounds of sums are omitted in this proof if they take the values of ±∞. Recall that the HBM equations can be expressed as the harmonic residual of the dynamics (29). Let (26) be a linear basis for the Koopman lift as in Theorem 2 with $\mathbf{z}(t) \approx \boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}}(\mathbf{y}(t), t)$. With the ordering (26), the first $2N_{\mathbf{u}} + 1$ entries of $\mathbf{z}$ approximate all basis functions connected to y_1, the next $2N_{\mathbf{u}} + 1$ entries approximate all basis functions connected to y_2 et cetera. With this separation, divide the matrix $\mathbf{B}(\mathbf{p}) \in \mathbb{C}^{n(2N_{\mathbf{u}}+1)\times(2N_{\mathbf{u}}+1)}$ into n square matrices $\{\mathbf{B}_l\}_{l=1}^{n}$ such that $\mathbf{B}_l \in \mathbb{C}^{(2N_{\mathbf{u}}+1)\times(2N_{\mathbf{u}}+1)}$ influences the derivative of y_l. Below, the argument is made for one $\mathbf{B}_l$ and $\mathbf{B}$ can be stacked together in the end.

Consider now the time derivative along the system flow of a Koopman basis function corresponding to the l-th state and frequency j, evaluated on the perturbed system:

$$\psi_{j,l}(\mathbf{y}, t) = y_l \mathrm{e}^{ij\omega t} \tag{B5}$$

$$= x_l \mathrm{e}^{ij\omega t} - \sum_k p_{k,l} \mathrm{e}^{i(j+k)\omega t} \tag{B6}$$

$$\dot{\psi}_{j,l}(\mathbf{y}, t) = \dot{x}_l \mathrm{e}^{ij\omega t} + ij\omega x_l \mathrm{e}^{ij\omega t} - \sum_k i(j+k)\omega p_{k,l} \mathrm{e}^{i(j+k)\omega t} \tag{B7}$$

$$= f_l(\mathbf{y} + \mathbf{x}_p, t)\mathrm{e}^{ij\omega t} + ij\omega y_l \mathrm{e}^{ij\omega t} - \sum_k ik\omega p_{k,l} \mathrm{e}^{i(j+k)\omega t}. \tag{B8}$$

The inner product $\langle \dot{\psi}_{j,l}, \boldsymbol{\Psi}_{\mathbf{z},\mathrm{lin}} \rangle$ is a row of $\mathbf{A}$ and $\langle \dot{\psi}_{j,l}, \mathbf{u} \rangle$ is a row of $\mathbf{B}_l$. As the inner product for $\mathbf{A}$ contains all summands that are dependent on $\mathbf{y}$ and the inner product for $\mathbf{B}$ collects all parts that are not dependent on $\mathbf{y}$, we can set $\mathbf{y} = \mathbf{0}$ in this proof about properties of $\mathbf{B}$. With this, it results

$$\dot{\psi}_{j,l} = f_l(\mathbf{x}_p(t), t)\mathrm{e}^{ij\omega t} - \left(\sum_k ik\omega p_{k,l} \mathrm{e}^{ik\omega t} \right) \mathrm{e}^{ij\omega t} \tag{B9}$$

$$= r_l(t)\mathrm{e}^{ij\omega t} \tag{B10}$$

$$= \sum_k r_{k,l} \mathrm{e}^{i(k+j)\omega t} \tag{B11}$$

$$= \sum_{\tilde{k}} r_{\tilde{k}-j,l} \mathrm{e}^{i\tilde{k}\omega t}. \tag{B12}$$

The inner product with $\mathbf{u}$ is just a projection onto the Fourier basis functions between $\pm N_{\mathbf{u}}$, and therefore, the row of $\mathbf{B}_l$ that corresponds to frequency j is given by

$$\langle \dot{\psi}_{j,l}, \mathbf{u} \rangle = (r_{-N_{\mathbf{u}}-j,l}, \ldots, r_{N_{\mathbf{u}}-j,l}). \tag{B13}$$

Comparing two successive rows, i.e., the j-th and the $(j+1)$-th row of $\mathbf{B}_l$, it is obvious from (B13) that they contain the same entries, but shifted to the right by one element. Adequately, going from row j to row $j+1$, the entry $r_{-N_{\mathbf{u}}-(j+1),l}$ is added on the left, while the entry $r_{N_{\mathbf{u}}-j,l}$ is pushed out on the right. The matrix $\mathbf{B}_l$ thus has a Toeplitz structure, where all residuals between $r_{-2N_{\mathbf{u}},l}$ (in the bottom left entry of $\mathbf{B}_l$ with $j = N_{\mathbf{u}}$) and $r_{2N_{\mathbf{u}},l}$ (in the top right entry with $j = -N_{\mathbf{u}}$) occur, yielding $4N_{\mathbf{u}} + 1$ distinct entries in total.

In particular, if the coefficients $\mathbf{p}$ describe the HBM solution for order $2N_{\mathbf{u}}$, then all these residuals vanish and it holds that $\mathbf{B}(\mathbf{p}) = \mathbf{0}$. □

The arguments of the previous proof can be reused for Theorem 1 with slightly modified assumptions. Recall the theorem:

Theorem 1 *Let $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{p})\mathbf{z} + \mathbf{B}(\mathbf{p})\mathbf{u}$ be the lifted dynamics of frequency order N_{HBM} of system (6) around an unknown periodic ansatz of the form (10). The N_{HBM}-th order HBM equations (11), i.e., $\mathbf{r}_k = 0$, $|k| \le N_{\text{HBM}}$, are given by $\mathbf{C}_{\mathbf{z}}\mathbf{B}(\mathbf{p}) = \mathbf{0}$, where $\mathbf{C}_{\mathbf{z}}$ is the constant selection matrix that fulfills $\mathbf{y} = \mathbf{C}_{\mathbf{z}}\boldsymbol{\Psi}_{\mathbf{z}}(\mathbf{y}, t)$ for all t.*

Proof In contrast to the previously proven Theorem 2, in this theorem the basis functions are not restricted to be linear in the state. Due to sesquilinearity of the inner product (2b), however, the selection matrix $\mathbf{C}_{\mathbf{z}}$ can be pulled into the expression for $\mathbf{B}(\mathbf{p})$ via

$$\mathbf{C}_{\mathbf{z}}\mathbf{B}(\mathbf{p}) = \mathbf{C}_{\mathbf{z}} \langle \dot{\boldsymbol{\Psi}}_{\mathbf{z}}, \mathbf{u} \rangle \tag{B14}$$

$$= \langle \mathbf{C}_{\mathbf{z}}\dot{\boldsymbol{\Psi}}_{\mathbf{z}}, \mathbf{u} \rangle \tag{B15}$$

$$= \left\langle \begin{pmatrix} \dot{\psi}_{0,1} \\ \vdots \\ \dot{\psi}_{0,n} \end{pmatrix}, \mathbf{u} \right\rangle. \tag{B16}$$

These expressions can again be evaluated row-wise with (B13) and since $j = 0$ for all entries of (B14), the condition $\mathbf{C}_{\mathbf{z}}\mathbf{B}(\mathbf{p}) = 0$ is equivalent to $\mathbf{r}_k = 0$ for all $|k| \le N_{\text{HBM}}$, which are exactly the HBM equations. □

B.2 Theorem 3

Similarly to the theorems proven previously, it is central in the proof for Theorem 3 to evaluate the inner product explicitly. The theorem is restated for convenience.

Theorem 3 *Let $\dot{\mathbf{z}} = \mathbf{A}(\mathbf{p})\mathbf{z} + \mathbf{B}(\mathbf{p})\mathbf{u}$ be the lifted dynamics around a periodic solution of system (6) with linear basis functions $\boldsymbol{\Psi}_{\mathbf{z},\text{lin}}$ of frequency order $N_{\mathbf{u}}$ that are ordered as in (22). Then, the Hill matrix $\mathbf{H}$, truncated to frequency order $N_{\mathbf{u}}$, for the periodic solution parameterized by $\mathbf{p}$ results from the matrix $\mathbf{A}(\mathbf{p})$ by the similarity transform $\mathbf{H} = \mathbf{U}\mathbf{A}(\mathbf{p})\mathbf{U}^{\mathrm{T}}$, where $\mathbf{U}$ is an orthogonal permutation matrix that satisfies $\mathbf{U}\boldsymbol{\Psi}_{\mathbf{z},\text{lin}} = (\mathbf{y}^{\mathrm{T}}\,\mathrm{e}^{iN_{\mathbf{u}}\omega t}, \ldots, \mathbf{y}^{\mathrm{T}}\,\mathrm{e}^{-iN_{\mathbf{u}}\omega t})^{\mathrm{T}}$.*

Proof The matrix $\mathbf{U}$ is derived from the identity matrix by reordering its rows, and it reorders the entries of $\boldsymbol{\Psi}_{\mathbf{z},\text{lin}}$ such that the resulting vector has entries descending in frequency, where all states corresponding to the same frequency are collected together. Denote $\tilde{\boldsymbol{\Psi}}_{\mathbf{z}} := \mathbf{U}\boldsymbol{\Psi}_{\mathbf{z},\text{lin}} =: ((\tilde{\boldsymbol{\Psi}}_{\mathbf{z}})^{\mathrm{T}}_{-N_{\mathbf{u}}}, \ldots, (\tilde{\boldsymbol{\Psi}}_{\mathbf{z}})^{\mathrm{T}}_{N_{\mathbf{u}}})^{\mathrm{T}}$ and separate it into $2N_{\mathbf{u}}+1$ blocks of length n with $(\tilde{\boldsymbol{\Psi}}_{\mathbf{z}})_k := \mathbf{y}\mathrm{e}^{-ik\omega t}$. First, we show that the Koopman lift performed with basis $\tilde{\boldsymbol{\Psi}}_{\mathbf{z}}$ identically yields $\mathbf{H}$ as its system matrix $\tilde{\mathbf{A}}$, and afterward we demonstrate how performing the lift with $\boldsymbol{\Psi}_{\mathbf{z},\text{lin}}$ instead of $\tilde{\boldsymbol{\Psi}}_{\mathbf{z}}$ demands the additional similarity transform.

As in the proof for Theorem 2, we start by determining the time derivative of the basis functions. For the vector of time derivatives, by the chain rule it holds that

$$\dot{\tilde{\boldsymbol{\Psi}}}_{\mathbf{z}} = \frac{\partial \tilde{\boldsymbol{\Psi}}_{\mathbf{z}}}{\partial \mathbf{y}}\tilde{\mathbf{f}} + \frac{\partial \tilde{\boldsymbol{\Psi}}_{\mathbf{z}}}{\partial t}\ , \tag{B17}$$

where $\tilde{\mathbf{f}}$ is the nonlinear perturbed dynamics (27a). Evaluating the individual summands of (B17) yields

$$\left.\frac{\partial \tilde{\boldsymbol{\Psi}}_{\mathbf{z}}}{\partial t}\right|_{\mathbf{y},t} = \begin{pmatrix} iN_{\mathbf{u}}\omega\ \mathbf{y}\mathrm{e}^{iN_{\mathbf{u}}\omega t} \\ \vdots \\ -iN_{\mathbf{u}}\omega\ \mathbf{y}\mathrm{e}^{-iN_{\mathbf{u}}\omega t} \end{pmatrix} \tag{B18}$$

$$\left.\frac{\partial \tilde{\boldsymbol{\Psi}}_{\mathbf{z}}}{\partial \mathbf{y}}\right|_{\mathbf{y},t} = \begin{pmatrix} \mathbf{I}\mathrm{e}^{iN_{\mathbf{u}}\omega t} \\ \vdots \\ \mathbf{I}\mathrm{e}^{-iN_{\mathbf{u}}\omega t} \end{pmatrix} \tag{B19}$$

$$\tilde{\mathbf{f}}(\mathbf{y}, t) = \sum_{j=-\infty}^{\infty} \mathbf{J}_j \mathrm{e}^{ij\omega t}\mathbf{y} + \mathcal{O}(\|\mathbf{y}\|^2) + \mathbf{b}(t). \tag{B20}$$

The terms of order $\mathcal{O}(\|\mathbf{y}\|^2)$ in (B20) can be dropped because terms of higher polynomial order are not represented in the chosen basis $\boldsymbol{\Psi}_{\mathbf{z},\text{lin}}$. The terms in $\mathbf{b}(t)$ appear if the candidate ansatz is not a periodic solution. They are purely time-dependent and will remain so after multiplication by (B19). Hence they are collected in the $\mathbf{B}$ matrix. The column vector in (B18) and the matrix in (B19) can be decomposed into $2N_\mathbf{u}+1$ blocks with n rows and 1 or n columns as indicated above, each block corresponding to a specific frequency. These blocks are labeled in ascending order by k, $-N_\mathbf{u} \leq k \leq N_\mathbf{u}$, such that the k-th block corresponds to the term $\mathrm{e}^{-ik\omega t}$. Equations (B18)–(B20) are substituted into the k-th block in (B17) to yield

$$\left(\dot{\tilde{\boldsymbol{\Psi}}}_\mathbf{z}\right)_k = \mathrm{e}^{-ik\omega t}\left(\sum_{j=-\infty}^{\infty} \mathbf{J}_j \mathrm{e}^{ij\omega t}\mathbf{y}\right) - ik\omega\mathbf{y}\mathrm{e}^{-ik\omega t} \tag{B21}$$

$$= \sum_{j=-\infty}^{\infty} \mathbf{J}_j \mathrm{e}^{i(j-k)\omega t}\mathbf{y} - ik\omega\mathbf{y}\mathrm{e}^{-ik\omega t} \tag{B22}$$

$$= -ik\omega\left(\tilde{\boldsymbol{\Psi}}_\mathbf{z}\right)_k + \sum_{j=-\infty}^{\infty} \mathbf{J}_j\left(\tilde{\boldsymbol{\Psi}}_\mathbf{z}\right)_{k-j}\ , \tag{B23}$$

where the purely time-dependent terms and the terms of order $\mathcal{O}(\|\mathbf{y}\|^2)$ have been dropped for the sake of legibility. To obtain the k-th block of the Koopman lift matrix $\tilde{\mathbf{A}}$, the inner product $\left\langle\left(\dot{\tilde{\boldsymbol{\Psi}}}_\mathbf{z}\right)_k, \tilde{\boldsymbol{\Psi}}_\mathbf{z}\right\rangle$ is considered. Due to the sesquilinearity of the inner product, both summands in (B23) can be computed separately. It is easy to see that the first summand yields a sparse matrix where only the diagonal in the k-th column block is nonzero. With an index shift

$$\sum_{j=-\infty}^{\infty} \mathbf{J}_j\left(\tilde{\boldsymbol{\Psi}}_\mathbf{z}\right)_{k-j} = \sum_{\tilde{j}=-\infty}^{\infty} \mathbf{J}_{k-\tilde{j}}\left(\tilde{\boldsymbol{\Psi}}_\mathbf{z}\right)_{\tilde{j}}\ , \tag{B24}$$

the inner product with the second summand of (B23) yields the matrix

$$\left(\mathbf{J}_{k-(-N_\mathbf{u})} \ \dots \ \mathbf{J}_{k-N_\mathbf{u}}\right). \tag{B25}$$

Collecting all row blocks and both summands together, the Koopman lift matrix $\tilde{\mathbf{A}}$ for the re-sorted basis is given by

$$\tilde{\mathbf{A}} = \left\langle\dot{\tilde{\boldsymbol{\Psi}}}_\mathbf{z}, \tilde{\boldsymbol{\Psi}}_\mathbf{z}\right\rangle = \begin{pmatrix} \mathbf{J}_0 + iN_\mathbf{u}\omega & \dots & \mathbf{J}_{-2N_\mathbf{u}} \\ \vdots & \ddots & \vdots \\ \mathbf{J}_{2N_\mathbf{u}} & \dots & \mathbf{J}_0 - iN_\mathbf{u}\omega \end{pmatrix}. \tag{B26}$$

Comparison of (B26) to the truncated Hill matrix (18) shows the identity $\tilde{\mathbf{A}} = \mathbf{H}$ for the Koopman lift with the re-sorted basis.

Finally, the similarity transform obtained by the reordering is considered. The permutation matrix $\mathbf{U}$ is constant and can thus be pulled out of the time derivative. With the sesquilinearity of the matrix-valued inner product (2b), (2c), it then follows

$$\mathbf{H} = \tilde{\mathbf{A}} = \left\langle\dot{\tilde{\boldsymbol{\Psi}}}_\mathbf{z}, \boldsymbol{\Psi}_\mathbf{z}\right\rangle \tag{B27}$$

$$= \left\langle\mathbf{U}\dot{\boldsymbol{\Psi}}_{\mathbf{z},\text{lin}}, \mathbf{U}\boldsymbol{\Psi}_{\mathbf{z},\text{lin}}\right\rangle \tag{B28}$$

$$= \mathbf{U}\left\langle\dot{\boldsymbol{\Psi}}_{\mathbf{z},\text{lin}}, \boldsymbol{\Psi}_{\mathbf{z},\text{lin}}\right\rangle\mathbf{U}^* \tag{B29}$$

$$= \mathbf{U}\mathbf{A}(\mathbf{p})\mathbf{U}^\mathrm{T}\ , \tag{B30}$$

where $\mathbf{A}(\mathbf{p})$ is the Koopman lift with the standard basis. In the last step, it has been exploited that $\mathbf{U}$ is a real-valued matrix and thus transpose and conjugate transpose coincide. □

B.3 Symmetry properties of the optimization problem

Below, it will be shown that there always exist solutions to the optimization problem (56) fulfilling (41b), even if this constraint is not enforced explicitly. This is summarized in the following proposition.

Proposition 4 *For an arbitrary collection* $\hat{\mathbf{c}}_\text{all} \in \mathbb{C}^{n(2N_\mathbf{u}+1)}$ *of coefficients which admit the residual* $\mathbf{R}$ *in the variational equation (45), there exists another collection* $\hat{\mathbf{c}}_\text{sym} \in \mathbb{C}^{n(2N_\mathbf{u}+1)}$ *constructed by*

$$\hat{c}_{\text{sym},l,k} = \frac{1}{2}\left(\hat{c}_{l,k} + \bar{\hat{c}}_{l,-k}\right)\ , \tag{B31}$$

which fulfills (41b). The residual $\mathbf{R}_\text{sym}$ *of the variational equation (45) for* $\hat{\mathbf{c}}_\text{sym}$ *is real for arbitrary t with* $\mathbf{R}_\text{sym} = \mathrm{Re}(\mathbf{R})$.

The proof for this proposition is only sketched here. Some mathematical steps are only indicated but not performed in detail for the sake of brevity.

Proof Below, vectors $\mathbf{d} \in \mathbb{C}^{(2N_\mathbf{u}+1)}$ which fulfill (41b) will be called symmetric. The symmetric vectors form a vector space over $\mathbb{R}$ (but not over $\mathbb{C}$).

First, the structure of the Koopman lift matrix $\mathbf{A}$ is revisited. Applying the similarity transform of Theorem 3 to the Hill matrix explicitly yields for the matrix $\mathbf{A}$ the structure

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \dots & \mathbf{A}_{1n} \\ \vdots & \ddots & \vdots \\ \mathbf{A}_{n1} & \dots & \mathbf{A}_{nn} \end{pmatrix} \tag{B32}$$

$$\mathbf{A}_{lj} = \begin{pmatrix} J_{lj,0} & \dots & J_{lj,2N_\mathbf{u}} \\ \vdots & \ddots & \vdots \\ J_{lj,-2N_\mathbf{u}} & \dots & J_{lj,0} \end{pmatrix} + \delta_{lj}\mathrm{diag}(-i\omega N_\mathbf{u}, \dots, i\omega N_\mathbf{u}). \tag{B33}$$

Hence, the lj-th block of the matrix $\mathbf{A}$ contains the Fourier coefficients of the lj-th entry of the system matrix $\mathbf{J}$. In particular, since $\mathbf{J}(t)$ admits only real values, $J_{lj,k} = \overline{J_{lj,-k}}$ and each block $\mathbf{A}_{lj}$ fulfills a variant of the symmetry property: The $-k$-th row is the complex conjugate of the k-th row, with the order of elements reversed. A little algebra shows that for matrices with these properties and symmetric vectors $\mathbf{d} \in \mathbb{C}^{(2N_\mathbf{u}+1)}$, the symmetry is retained though multiplication: $\mathbf{A}_{lj}\mathbf{d}$ is symmetric if $\mathbf{d}$ is symmetric.

If a matrix $\mathbf{P} = [\mathbf{d}_{lj}]$ is constructed block-wise from symmetric column vectors $\{\mathbf{d}_{lj}\}_{l,j=1}^{n}$, then the matrix $\mathbf{AP}$ is again constructed from symmetric column vectors since all block entries of $\mathbf{AP}$ can be decoupled into sums of products $\mathbf{A}_{li}\mathbf{d}_{ij}$, which each retain the symmetry as described above. In particular, since the matrix $\mathbf{W}$ in (33) is constructed in the considered fashion, the product $\mathbf{AW}$ again has block-wise symmetric entries. Iterative application of this multiplicative invariance to

$$\mathbf{Q} = \mathrm{e}^{\mathbf{A}T}\mathbf{W} = \sum_{k=0}^{\infty} T^k \mathbf{A}^k \mathbf{W} \tag{B34}$$

$$\mathbf{L} = (\mathbf{A} + \mathbf{D})\mathbf{Q} \tag{B35}$$

shows that the column vectors $\mathbf{Q}_{lj}$ and $\mathbf{L}_{lj}$ in (50) are also symmetric.

Next, the scalar product $\mathbf{b}^\mathrm{T}\mathbf{d} \in \mathbb{C}$ of a symmetric vector $\mathbf{d}$ and an arbitrary vector $\mathbf{b}$ of appropriate size is considered. The operator flip $\mathbf{b}$ reverses the order of entries of a vector $\mathbf{b}$. Element-wise evaluation shows that

$$\left(\mathrm{flip}\,\overline{\mathbf{b}}\right)^\mathrm{T}\mathbf{d} = \sum_{k=-N_\mathbf{u}}^{N_\mathbf{u}} \overline{b}_{-k} d_k \tag{B36a}$$

$$= \sum_{k=-N_\mathbf{u}}^{N_\mathbf{u}} \overline{b}_k \overline{d}_k \tag{B36b}$$

$$= \overline{\left(\mathbf{b}^\mathrm{T}\mathbf{d}\right)}. \tag{B36c}$$

Defining $\hat{\mathbf{c}}_{l,\mathrm{sym}} = \frac{1}{2}\left(\hat{\mathbf{c}}_l + \mathrm{flip}\,\overline{\hat{\mathbf{c}}}_l\right)$ and using the above relation yields

$$\hat{\mathbf{c}}_{l,\mathrm{sym}}\mathbf{Q}_{ij} = \mathrm{Re}(\hat{\mathbf{c}}_l\mathbf{Q}_{ij}) , \tag{B37}$$

and analogously for $\mathbf{L}_{ij}$. As $\mathbf{R}$ is constructed from summands of this form via (49)–(52), the proposition follows. □

A purely real matrix $\mathbf{R}_\mathrm{sym}$ will necessarily have a Frobenius norm smaller or equal than the norm of a matrix with same real part and arbitrary imaginary part. Hence, Proposition 4 allows to easily construct a minimizer that satisfies all constraints in case a different minimizer is returned in the quadratic program (56).

References

1. Mauroy, A., Mezić, I., Susuki, Y. (eds.): The Koopman Operator in Systems and Control: Concepts, Methodologies and Applications. Lecture Notes in Control and Information Sciences, vol. 484. Springer, Cham (2020)
2. Brunton, S.L., Budišić, M., Kaiser, E., Kutz, J.N.: Modern Koopman theory for dynamical systems. SIAM Rev. **64**(2), 229–340 (2022). https://doi.org/10.1137/21m1401243
3. Williams, M.O., Kevrekidis, I.G., Rowley, C.W.: A data-driven approximation of the Koopman operator: extending dynamic mode decomposition. J. Nonlinear Sci. **25**(6), 1307–1346 (2015). https://doi.org/10.1007/s00332-015-9258-5
4. Kono, Y., Susuki, Y., Hikihara, T.: Modeling of advective heat transfer in a practical building atrium via Koopman mode decomposition. In: Mauroy, A., Mezić, I., Susuki, Y. (eds.) The Koopman Operator in Systems and Control: Concepts, Methodologies, and Applications, pp. 481–506. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-35713-9_18
5. Korda, M., Mezić, I.: Koopman model predictive control of nonlinear dynamical systems. In: Mauroy, A., Mezić, I., Susuki, Y. (eds.) The Koopman Operator in Systems and Control: Concepts, Methodologies and Applications, pp. 235–255. Springer, Cham (2020)

6. Koopman, B.O.: Hamiltonian systems and transformation in Hilbert space. Proc. Natl. Acad. Sci. USA **17**(16577368), 315–318 (1931). https://doi.org/10.1073/pnas.17.5.315
7. Mezić, I.: Spectral properties of dynamical systems, model reduction and decompositions. Nonlinear Dyn. **41**(1), 309–325 (2005). https://doi.org/10.1007/s11071-005-2824-x
8. Mauroy, A., Mezić, I., Moehlis, J.: Isostables, isochrons, and Koopman spectrum for the action-angle representation of stable fixed point dynamics. Physica D **261**, 19–30 (2013). https://doi.org/10.1016/j.physd.2013.06.004
9. Mohr, R., Mezić, I.: Construction of eigenfunctions for scalar-type operators via Laplace averages with connections to the Koopman operator (2014). https://doi.org/10.48550/arXiv.1403.6559
10. Mauroy, A., Susuki, Y., Mezić, I.: Introduction to the Koopman operator in dynamical systems and control theory. In: Mauroy, A., Susuki, Y., Mezić, I. (eds.) The Koopman Operator in Systems and Control: Concepts, Methodologies and Applications, pp. 3–33. Springer, Cham (2020)
11. Nayfeh, A.H., Mook, D.T., Holmes, P.: Nonlinear oscillations. J. Appl. Mech. **47**(3), 692–692 (1980). https://doi.org/10.1115/1.3153771
12. Bentvelsen, B., Lazarus, A.: Modal and stability analysis of structures in periodic elastic states: application to the Ziegler column. Nonlinear Dyn. **91**(2), 1349–1370 (2018). https://doi.org/10.1007/s11071-017-3949-4
13. Karkar, S., Vergez, C., Cochelin, B.: Oscillation threshold of a clarinet model: a numerical continuation approach. J. Acoust. Soc. Am. **131**(1), 698–707 (2012). https://doi.org/10.1121/1.3651231
14. Noiray, N., Durox, D., Schuller, T., Candel, S.: A unified framework for nonlinear combustion instability analysis based on the flame describing function. J. Fluid Mech. **615**, 139–167 (2008). https://doi.org/10.1017/S0022112008003613
15. da Cruz Scarabello, M., Messias, M.: Bifurcations leading to nonlinear oscillations in a 3D piecewise linear memristor oscillator. Int. J. Bifurc. Chaos **24**(01), 1430001 (2014). https://doi.org/10.1142/s0218127414300018
16. Cheffer, A., Savi, M.A., Pereira, T.L., de Paula, A.S.: Heart rhythm analysis using a nonlinear dynamics perspective. Appl. Math. Model. **96**, 152–176 (2021). https://doi.org/10.1016/j.apm.2021.03.014
17. Peeters, M., Viguié, R., Sérandour, G., Kerschen, G., Golinval, J.-C.: Nonlinear normal modes, Part II: toward a practical computation using numerical continuation techniques. Mech. Syst. Signal Process. **23**(1), 195–216 (2009). https://doi.org/10.1016/j.ymssp.2008.04.003
18. Kevrekidis, I.G., Aris, R., Schmidt, L.D., Pelikan, S.: Numerical computation of invariant circles of maps. Physica D **16**(2), 243–251 (1985). https://doi.org/10.1016/0167-2789(85)90061-2
19. Parker, T.S., Chua, L.O.: Practical Numerical Algorithms for Chaotic Systems. Springer, Berlin (1989)
20. Morrison, D.D., Riley, J.D., Zancanaro, J.F.: Multiple shooting method for two-point boundary value problems. Commun. ACM **5**(12), 613–614 (1962). https://doi.org/10.1145/355580.369128
21. Ascher, U.M., Mattheij, R.M.M., Russell, R.D.: Numerical Solution of Boundary Value Problems for Ordinary Differential Equations. Soc. Ind. Appl. Math. (1995). https://doi.org/10.1137/1.9781611971231
22. Krack, M., Gross, J.: Harmonic Balance for Nonlinear Vibration Problems. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-14023-6
23. Detroux, T., Renson, L., Masset, L., Kerschen, G.: The harmonic balance method for bifurcation analysis of large-scale nonlinear mechanical systems. Comput. Methods Appl. Mech. Eng. **296**, 18–38 (2015). https://doi.org/10.1016/j.cma.2015.07.017
24. Cameron, T.M., Griffin, J.H.: An alternating frequency/time domain method for calculating the steady-state response of nonlinear dynamic systems. J. Appl. Mech. **56**(1), 149–154 (1989). https://doi.org/10.1115/1.3176036
25. Cochelin, B., Vergez, C.: A high order purely frequency-based harmonic balance formulation for continuation of periodic solutions. J. Sound Vib. **324**(1–2), 243–262 (2009). https://doi.org/10.1016/j.jsv.2009.01.054
26. Zhou, J., Hagiwara, T., Araki, M.: Spectral characteristics and eigenvalues computation of the harmonic state operators in continuous-time periodic systems. Syst. Control Lett. **53**(2), 141–155 (2004). https://doi.org/10.1016/j.sysconle.2004.03.002
27. Guillot, L., Lazarus, A., Thomas, O., Vergez, C., Cochelin, B.: A purely frequency based Floquet–Hill formulation for the efficient stability computation of periodic solutions of ordinary differential systems. J. Comput. Phys. **416**, 109477 (2020). https://doi.org/10.1016/j.jcp.2020.109477
28. Lazarus, A., Thomas, O.: A harmonic-based method for computing the stability of periodic solutions of dynamical systems. Comptes Rendus Mécanique **338**(9), 510–517 (2010). https://doi.org/10.1016/j.crme.2010.07.020
29. Peletan, L., Baguet, S., Torkhani, M., Jacquet-Richardet, G.: A comparison of stability computational methods for periodic solution of nonlinear problems with application to rotordynamics. Nonlinear Dyn. **72**(3), 671–682 (2013). https://doi.org/10.1007/s11071-012-0744-0
30. Moore, G.: Floquet theory as a computational tool. SIAM J. Numer. Anal. **42**(6), 2522–2568 (2005). https://doi.org/10.1137/s0036142903434175
31. Wu, J., Hong, L., Jiang, J.: A robust and efficient stability analysis of periodic solutions based on harmonic balance method and Floquet–Hill formulation. Mech. Syst. Signal Process. **173**, 109057 (2022). https://doi.org/10.1016/j.ymssp.2022.109057
32. Bayer, F., Leine, R.I.: A Koopman view on the harmonic balance and Hill method. In: Proceedings of the 10th European Nonlinear Dynamics Conference (2022). https://enoc2020.sciencesconf.org/394116
33. Naylor, A.W., Sell, G.R.: Linear Operator Theory in Engineering and Science. Holt, Rinehart & Winston, New York (1971)
34. Budišić, M., Mohr, R., Mezić, I.: Applied Koopmanism. Chaos: Interdiscip. J. Nonlinear Sci. **22**(4), 047510 (2012). https://doi.org/10.1063/1.4772195
35. Carleman, T.: Application de la théorie des équations intégrales linéaires aux systèmes d'équations différentielles non linéaires. Acta Math. **59**, 63–87 (1932)
36. Berrueta, T.A., Abraham, I., Murphey, T.: Experimental applications of the Koopman operator in active learning for control. In: Mauroy, A., Mezić, I., Susuki, Y.

(eds.) The Koopman Operator in Systems and Control, pp. 421–450. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-35713-9_16

37. Champion, K.P., Brunton, S.L., Kutz, J.N.: Discovery of nonlinear multiscale systems: sampling strategies and embeddings. SIAM J. Appl. Dyn. Syst. **18**(1), 312–333 (2019). https://doi.org/10.1137/18m1188227
38. Bittanti, S., Colaneri, P.: Invariant representations of discrete-time periodic systems. Automatica **36**(12), 1777–1793 (2000). https://doi.org/10.1016/S0005-1098(00)00087-X
39. Teschl, G.: Ordinary Differential Equations and Dynamical Systems. Graduate Studies in Mathematics, vol. 140. American Mathematical Society, Providence, Rhode Island (2012)
40. Von Groll, G., Ewins, D.J.: The harmonic balance method with arc-length continuation in rotor/stator contact problems. J. Sound Vib. **241**(2), 223–233 (2001). https://doi.org/10.1006/jsvi.2000.3298
41. Golub, G.H.V., Van Loan, C.F.V.: Matrix Computations. North Oxford Academic, Oxford (1986)
42. Bruder, D., Fu, X., Gillespie, R.B., Remy, C.D., Vasudevan, R.: Data-driven control of soft robots using Koopman operator theory. IEEE Trans. Rob. **37**(3), 948–961 (2021). https://doi.org/10.1109/TRO.2020.3038693
43. Johnson, C.A., Balakrishnan, S., Yeung, E.: Heterogeneous mixtures of dictionary functions to approximate subspace invariance in Koopman operators (2022). https://doi.org/10.48550/arXiv.2206.13585
44. Korda, M., Mezić, I.: Linear predictors for nonlinear dynamical systems: Koopman operator meets model predictive control. Automatica **93**, 149–160 (2018). https://doi.org/10.1016/j.automatica.2018.03.046
45. Moler, C., Van Loan, C.: Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. SIAM Rev. **45**(1), 3–49 (2003). https://doi.org/10.1137/S00361445024180
46. Al-Mohy, A.H., Higham, N.J.: Computing the action of the matrix exponential, with an application to exponential integrators. SIAM J. Sci. Comput. **33**(2), 488–511 (2011). https://doi.org/10.1137/100788860
47. Magnus, W., Winkler, S.: Hill's Equation. Interscience Publishers, New York (1966)
48. Moideen, H., Falzarano, J., Somayajula, A.: Parametric roll of container ships in head waves. Ocean Systems Engineering 2: 239–255. (2012). https://doi.org/10.12989/ose.2012.2.4.239 https://doi.org/10.12989/ose.2012.2.4.239
49. Leine, R.I.: Non-smooth stability analysis of the parametrically excited impact oscillator. Int. J. Non-Linear Mech. **47**(9), 1020–1032 (2012). https://doi.org/10.1016/j.ijnonlinmec.2012.06.010
50. Kovacic, I., Rand, R., Mohamed Sah, S.: Mathieu's equation and its generalizations: overview of stability charts and their features. Appl. Mech. Rev. **70**(2) (2018). https://doi.org/10.1115/1.4039144
51. Abboud, E., Thomas, O., Grolet, A., Mahé, H.: On the solution of the Mathieu equation with multiple harmonic stiffness, parametric amplification for constant and harmonic forcing. In: Proceedings of the 10th European Nonlinear Dynamics Conference (2022). https://enoc2020.sciencesconf.org/341631
52. Bartuccelli, M.V., Gentile, G., Georgiou, K.V.: On the stability of the upside-down pendulum with damping. Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences **458**(2018), 255–269 (2001). https://doi.org/10.1098/rspa.2001.0859
53. Schiele, K., Hemmecke, R.: Migration effects in driven multiple pendula. Z. Angew. Math. Mech. **81**(5), 291–303 (2001)
54. Meirovich, L.: Methods of Analytical Dynamics. Advanced Engineering Series, vol. 3. McGraw-Hill Book Company, New York (1970)
55. Higham, N.: Matrix exponential times a vector. MATLAB Central File exchange (2022). www.mathworks.com/matlabcentral/fileexchange/29576-matrix-exponential-times-a-vector
56. Debeurre, M., Grolet, A., Mattei, P.-O., Cochelin, B., Thomas, O.: Nonlinear modes of cantilever beams at extreme amplitudes: numerical computation and experiments. In: Brake, M.R.W., Renson, L., Kuether, R.J., Tiso, P. (eds.) Nonlinear Struct. Syst., vol. 1, pp. 245–248. Springer, Cham (2023)
57. Hubay, C.Á., Kalmár-Nagy, T.: Return time approximation in planar nonlinear systems. J. Sound Vib. **508**, 116200 (2021). https://doi.org/10.1016/j.jsv.2021.116200
58. Hubay, C.Á., Kalmár-Nagy, T.: Period approximation for nonlinear oscillators with Carleman linearization. In: Proceedings of the 10th European Nonlinear Dynamics Conference (2022). https://enoc2020.sciencesconf.org/309588
59. Reed, M., Simon, B.: Functional Analysis. Methods of Modern Mathematical Physics, vol. 1, p. 400. Academic Press, San Diego (1980)
60. Ehrenfest, P., Kamerlingh Onnes, H.: XXXIII. Simplified deduction of the formula from the theory of combinations which Planck uses as the basis of his radiation theory. Lond Edinb Dublin Philos Mag J Sci **29**(170), 297–301 (1915). https://doi.org/10.1080/14786440208635308