MOX-Report No. 08/2026

Newmark time marching as a preconditioned iteration for large SPD linear systems

Micheletti, S.

MOX, Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

mox-dmat@polimi.it https://mox.polimi.it

Newmark time marching as a preconditioned iteration for large SPD linear systems

S. Micheletti

January 20, 2026

MOX - Dipartimento di Matematica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133 Milano, IT, stefano.micheletti@polimi.it

Abstract

We revisit the classical idea of computing the solution of a large symmetric positive definite linear system as the steady state of an artificial transient dynamics. Starting from a second-order model with "mass" and "damping" operators, we show that the explicit Newmark scheme ($\beta = 0$) induces a family of stationary iterations for $A\mathbf{u} = \mathbf{b}$, with a natural residual form and a transparent role for the pair (M, C). This viewpoint unifies, within a single algebraic framework, time-marching solvers, preconditioned fixed-point iterations, and momentum-like recurrences. In particular, by choosing M as a preconditioner and using Rayleigh damping $C = a_0 M + a_1 A$, the induced iteration can be implemented efficiently via inexact inner solves and behaves as a robust preconditioned method in challenging diffusion problems. Numerical experiments on model elliptic operators illustrate the influence of (M, C) and provide practical parameter guidelines.

Keywords: Newmark method; stationary iterations; preconditioning; Rayleigh damping; conjugate gradients; inexact solves.
MSC (2020): 65F10; 65N22; 65N30.

1 Introduction

We consider the linear system

$$A\,\overline{\mathbf{u}} = \mathbf{b}, \tag{1}$$

where $A \in \mathbb{R}^{n\times n}$ is symmetric positive definite, $\overline{\mathbf{u}} \in \mathbb{R}^n$ is the unknown solution vector, and $\mathbf{b} \in \mathbb{R}^n$ is the given right-hand side. System (1) admits a unique solution.

Throughout these notes we assume that n is very large, so that using a direct solver for (1) is prohibitively expensive. We therefore focus on iterative methods (see, e.g., [20] for

historical context). Our derivation will be somewhat unconventional: we will obtain iterations by discretizing the second-order initial value problem

$$\begin{cases} M\,\ddot{\mathbf{u}} + C\,\dot{\mathbf{u}} + A\,\mathbf{u} = \mathbf{b}, \quad t > 0 \\ \mathbf{u}(0) = \mathbf{u}^{(0)} \\ \dot{\mathbf{u}}(0) = \mathbf{v}^{(0)}. \end{cases} \tag{2}$$

System (2) can be interpreted as the dynamics of a virtual mechanical (or structural) mass–spring–damper system associated with (1): $M \in \mathbb{R}^{n\times n}$ and $C \in \mathbb{R}^{n\times n}$ are symmetric positive definite mass and damping matrices, while A plays the role of the stiffness matrix. The state $\mathbf{u} = \mathbf{u}(t)$ represents the "displacement" of the virtual system, and $\mathbf{u}^{(0)}$, $\mathbf{v}^{(0)}$ are the initial displacement and velocity. In particular, $\mathbf{u}^{(0)}$ plays the role of the initial guess in a standard iterative method for (1).

After introducing the continuous-in-time system (2), we turn to its time discretization. The main message of these notes is the following: *discretizing* (2) *can be viewed as an iterative process for solving* (1), as we show in Section 2.
Using transient dynamics to compute (or precondition) a static equilibrium problem is not a new idea: it is reminiscent of *dynamic relaxation* methods in structural mechanics [15, 16, 17]. In the nonlinear setting, closely related time-marching strategies appear under the name *pseudo-transient continuation* [18, 19].

1.1 Positioning and contributions

Using a transient dynamics to recover a static equilibrium is classical in structural mechanics: *dynamic relaxation* amounts to integrating an artificially damped vibration until a steady state is reached (see, e.g., Day's early note [15] and the subsequent developments by Otter–Cassell–Hobbs [16] and Underwood [17]). From the viewpoint of nonlinear solvers, pseudo-transient continuation plays an analogous role: a transient term is used as a globalization mechanism for Newton-like iterations, with the time step acting as a continuation parameter [18, 19]. In the linear SPD setting, it is also well known that *second-order* recurrences can be interpreted as polynomial accelerations of Richardson-type iterations (e.g. Chebyshev semi-iteration, second-order Richardson), and they are closely related to the "heavy-ball" (or momentum) method introduced by Polyak [4].

The goal of this note is not to compete with Krylov methods such as (P)CG as general-purpose black-box solvers for (1). Rather, we aim at a transparent bridge between a standard structural time integrator (Newmark) and a family of *preconditioned stationary iterations* whose building blocks are "mass" and "damping" operators. This viewpoint can be useful when (i) one wants iterations with a clear physical interpretation and tunable dissipation, (ii) one has efficient solvers for matrices built from M and C (e.g. sparse factorizations, multigrid smoothers, domain-decomposition blocks), or (iii) one wishes to embed the linear solve in a broader time-marching or nonlinear framework.

Concretely, the paper makes the following points.

- We show that the explicit Newmark scheme ($\beta = 0$) applied to (2) yields an iterative process for (1) in residual form, requiring only (a) one application of M^{-1}, (b) one multiplication by A per step to update the residual, and (c) one solve with $M + \gamma\tau_k C$ (Section 3).
- We identify a "clean" case in which proportional damping $C = \eta M$ reduces the Newmark-induced iteration to a preconditioned heavy-ball method, making the role of the parameters (τ_k, η) explicit (Section 3.4).
- We analyze commuting damping choices (in particular Rayleigh damping $C = a_0 M + a_1 A$) through a mode-wise three-term recurrence, derive contraction factors in terms of the generalized spectrum of (A, M), and propose a simple parameter-tuning strategy based on spectral bounds (Section 3.3–3.3.1 and Section 3.6).
- We illustrate, on model SPD problems (including a high-contrast diffusion benchmark), how the choice of M and C can turn the same Newmark time marching scheme from a "smoother-like" process into a robust solver, and we quantify the cost of the inexact Rayleigh variant via the total number of inner PCG steps (Section 6).

The paper is organized as follows. Section 2 recalls the Newmark discretization and motivates the explicit choice $\beta = 0$. Section 3 derives the induced iteration and its residual form. Sections 3.3–3.6 discuss the influence of M and C and provide design guidelines. Section 6 reports numerical experiments. Section 7 summarizes the main findings and outlines possible extensions, and the appendices collect stability proofs for the underlying continuous-time dynamics. Before delving into the main discussion, we anticipate a theoretical point that makes the correspondence meaningful. First, $(2)_1$ can be obtained from the following state equations:

$$\begin{cases} \mathbf{r} = \mathbf{b} - A\,\mathbf{u} & \text{definition of residual (external minus elastic force)} \\ \dot{\mathbf{u}} = \mathbf{v} & \text{definition of velocity} \\ \dot{\mathbf{r}} = -A\,\mathbf{v} & \text{consequence of } (3)_1 \text{ and } (3)_2 \\ M\,\dot{\mathbf{v}} = \mathbf{r} - C\,\mathbf{v} & \text{Newton's } 2^{\text{nd}} \text{ law with damping} \\ \mathbf{u}(0) = \mathbf{u}^{(0)} & \text{initial condition for } \mathbf{u} \\ \mathbf{v}(0) = \mathbf{v}^{(0)} & \text{initial condition for } \mathbf{v}. \end{cases} \tag{3}$$

The second-order equation $(2)_1$ follows by combining the relations in (3), i.e., by rewriting $(3)_4$ using $(3)_1$–$(3)_2$. Note that $\mathbf{u}(t) \equiv \overline{\mathbf{u}}$ is always an equilibrium (stationary) solution of $(2)_1$. Conversely, in order for (2) to be consistent with (1), we require the asymptotic constraint

$$\lim_{t\to+\infty} \mathbf{u}(t) = \overline{\mathbf{u}} \quad \forall \mathbf{u}^{(0)}, \mathbf{v}^{(0)} \in \mathbb{R}^n. \tag{4}$$

In the language of classical iterative methods for (1), this requirement plays the role of *convergence*, although it is stated at the continuous-time level. We will prove (4) in two different ways in Appendices A and B.

To this end, introduce the new variable $\mathbf{w} = \mathbf{u} - \overline{\mathbf{u}}$, which transforms (2) into the homogeneous system

$$\begin{cases} M\,\ddot{\mathbf{w}} + C\,\dot{\mathbf{w}} + A\,\mathbf{w} = 0, \quad t > 0 \\ \mathbf{w}(0) = \mathbf{w}^{(0)} \\ \dot{\mathbf{w}}(0) = \mathbf{w}_1^{(0)}, \end{cases} \tag{5}$$

where $\mathbf{w}^{(0)} = \mathbf{u}^{(0)} - \overline{\mathbf{u}}$ and $\mathbf{w}_1^{(0)} = \mathbf{v}^{(0)}$ are still arbitrary initial data. From the viewpoint of dynamical systems, (5) has an equilibrium point at the origin. Accordingly, instead of (4) we shall refer to

$$\lim_{t\to+\infty} \mathbf{w}(t) = 0 \quad \forall \mathbf{w}^{(0)}, \mathbf{w}_1^{(0)} \in \mathbb{R}^n. \tag{6}$$

Property (6) is equivalent to global asymptotic stability (in the sense of Lyapunov) of (5). The first proof, based on a modal analysis, requires additional assumptions on the damping matrix; the second proof is more general and does not introduce any extra constraint. In the remainder of the paper we assume that (6) (and hence (4)) holds.

2 Numerical discretization

We now discretize (2) in time using the well-known Newmark method [1, 2]. This is a one-step marching scheme for the displacement–velocity pair $(\mathbf{u}^{(k)}, \mathbf{v}^{(k)})$, where $\mathbf{u}^{(k)} \simeq \mathbf{u}(t_k)$ and $\mathbf{v}^{(k)} \simeq \dot{\mathbf{u}}(t_k)$, on a (possibly non-uniform) time grid $\{t_k\}_{k\geq 0}$ with $t_{k+1} = t_k + \tau_k$. For $k \geq 0$, Newmark's method reads as follows: find $(\mathbf{u}^{(k+1)}, \mathbf{v}^{(k+1)})$ such that

$$\begin{cases} M\,\mathbf{u}^{(k+1)} &=& M\,\mathbf{u}^{(k)} + \tau_k\,M\,\mathbf{v}^{(k)} \\ &+& \tau_k^2\big(\beta\,(\mathbf{b} - A\,\mathbf{u}^{(k+1)} - C\,\mathbf{v}^{(k+1)}) + (\tfrac{1}{2} - \beta)(\mathbf{b} - A\,\mathbf{u}^{(k)} - C\,\mathbf{v}^{(k)})\big) \\ M\,\mathbf{v}^{(k+1)} &=& M\,\mathbf{v}^{(k)} + \tau_k\big(\gamma(\mathbf{b} - A\,\mathbf{u}^{(k+1)} - C\,\mathbf{v}^{(k+1)}) + (1-\gamma)(\mathbf{b} - A\,\mathbf{u}^{(k)} - C\,\mathbf{v}^{(k)})\big), \end{cases} \tag{7}$$

where β, γ are nonnegative parameters. In block form, (7) can be rewritten as

$$\begin{aligned} &\begin{pmatrix} M + \beta\,\tau_k^2\,A & \beta\,\tau_k^2\,C \\ \gamma\,\tau_k\,A & M + \gamma\,\tau_k\,C \end{pmatrix} \begin{pmatrix} \mathbf{u}^{(k+1)} \\ \mathbf{v}^{(k+1)} \end{pmatrix} \\ &= \begin{pmatrix} M - (\frac{1}{2} - \beta)\,\tau_k^2\,A & \tau_k\,M - (\frac{1}{2} - \beta)\,\tau_k^2\,C \\ -(1-\gamma)\,\tau_k\,A & M - (1-\gamma)\,\tau_k\,C \end{pmatrix} \begin{pmatrix} \mathbf{u}^{(k)} \\ \mathbf{v}^{(k)} \end{pmatrix} + \begin{pmatrix} \frac{1}{2}\,\tau_k^2\,\mathbf{b} \\ \tau_k\,\mathbf{b} \end{pmatrix}. \end{aligned} \tag{8}$$

For $\beta > 0$ the scheme is implicit. When $\beta = 0$ it becomes explicit and can be solved by block forward substitution, provided one can afford solving linear systems whose coefficient matrix is a linear combination of M and C.

Thus, if we want to avoid treating the stiffness matrix A implicitly (which would require repeatedly solving (8) and could be even more expensive than solving (1) directly), we are naturally led to explicit methods. Alternatively, one could replace A in the first block row on the left-hand side of (8) by an approximation P (e.g., a preconditioner for A) and still obtain an implicit method. For this approach to be useful, however, solving systems with the block coefficient matrix

$$\begin{pmatrix} M + \beta\,\tau_k^2\,P & \beta\,\tau_k^2\,C \\ \gamma\,\tau_k\,A & M + \gamma\,\tau_k\,C \end{pmatrix}$$

must be substantially cheaper than applying a direct solver to (1).

Because this comparison is difficult to assess a priori, we start by studying explicit schemes and therefore set $\beta = 0$ from now on. Even in this case we will see that, with a suitable choice of the parameters τ_k and, in particular, of the matrices M and C, the resulting approach can be interpreted as a preconditioned iterative method applied to (1). The implicit case $\beta \neq 0$ is left for future work. A further and potentially impactful direction is to exploit the fact that Newmark's method naturally allows a variable time step τ_k (see (7)), whereas all experiments in this paper use a constant step size. Designing nonstationary choices $\{\tau_k\}$—for instance driven by online spectral estimates or by residual reduction—could turn the induced iteration into a Chebyshev-like polynomial process (cf. Remark 3.8) and may substantially reduce the outer iteration count on strongly ill-conditioned problems such as T2. We leave the systematic design, analysis, and cost assessment of adaptive step-size strategies to future work.

3 From Newmark time marching to preconditioned iterations

In this section we make the link between the explicit Newmark discretization ($\beta = 0$) and classical iterative solvers for (1) more explicit.

3.1 A residual form of the $\beta = 0$ scheme

Introduce the residual at step k,

$$\mathbf{r}^{(k)} = \mathbf{b} - A\,\mathbf{u}^{(k)}. \tag{9}$$

Setting $\beta = 0$ in (7) gives the update for the displacement,

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \tau_k\,\mathbf{v}^{(k)} + \frac{\tau_k^2}{2}\,M^{-1}\big(\mathbf{r}^{(k)} - C\,\mathbf{v}^{(k)}\big), \tag{10}$$

followed by the update for the velocity,

$$(M + \gamma\,\tau_k\,C)\,\mathbf{v}^{(k+1)} = (M - (1-\gamma)\,\tau_k\,C)\,\mathbf{v}^{(k)} + \tau_k\big(\gamma\,\mathbf{r}^{(k+1)} + (1-\gamma)\,\mathbf{r}^{(k)}\big). \tag{11}$$

Thus, each time step requires only: (i) one application of M^{-1} (i.e., one solve with M) in (10), (ii) one multiplication by A to form $\mathbf{r}^{(k+1)} = \mathbf{b} - A\mathbf{u}^{(k+1)}$, and (iii) one solve with $M + \gamma\tau_k C$. In particular, A never appears in a linear solve, which is what makes the scheme attractive as an "outer" iteration for (1).

Remark 3.1 (A dimensionally consistent CG-like update from a damped first-order dynamics)**.** *To connect with the classical conjugate-gradient notation, consider variables* $(\mathbf{u}, \mathbf{r}, \mathbf{v})$ *satisfying* $\mathbf{r} = \mathbf{b} - A\mathbf{u}$ *and interpret* $\mathbf{v} = \dot{\mathbf{u}}$ *as a* velocity*. In this mechanical viewpoint,* $\mathbf{r}$ *has the units of a force, while* $\mathbf{v}$ *has the units of a velocity, so the common algebraic normalization* $\mathbf{v}^{(k+1)} = \mathbf{r}^{(k+1)} + \beta_{k+1}^{\mathrm{CG}}\mathbf{v}^{(k)}$ *is not dimensionally meaningful. A consistent alternative is obtained by introducing a* scalar mass $m > 0$ *and a damping coefficient* $\delta \geq 0$ *and considering the first-order system*

$$\dot{\mathbf{u}} = \mathbf{v}, \qquad \dot{\mathbf{r}} = -A\mathbf{v}, \qquad m\,\dot{\mathbf{v}} = \mathbf{r} - \delta\,\mathbf{v},$$

where $\mathbf{r} = \mathbf{b} - A\mathbf{u}$*. With a step size* $\alpha_k = t_{k+1} - t_k$*, a forward Euler discretization of the first two equations (with* $\mathbf{v}$ *frozen over the step) gives* $\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \alpha_k\mathbf{v}^{(k)}$ *and* $\mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k A\mathbf{v}^{(k)}$*. Discretizing the damping law using the updated residual* $\mathbf{r}^{(k+1)}$ *yields*

$$\mathbf{v}^{(k+1)} = \mathbf{v}^{(k)} + \frac{\alpha_k}{m}\left(\mathbf{r}^{(k+1)} - \delta_k\mathbf{v}^{(k)}\right) = \beta_{k+1}^{\mathrm{CG}}\mathbf{v}^{(k)} + \rho_{k+1}\mathbf{r}^{(k+1)}, \qquad \rho_{k+1} = \frac{\alpha_k}{m}, \quad \beta_{k+1}^{\mathrm{CG}} = 1 - \frac{\alpha_k\delta_k}{m}.$$

Using backward Euler for the damping law, $m(\mathbf{v}^{(k+1)} - \mathbf{v}^{(k)})/\alpha_k = \mathbf{r}^{(k+1)} - \delta_k\,\mathbf{v}^{(k+1)}$*, leads to the same CG-like form* $\mathbf{v}^{(k+1)} = \beta_{k+1}^{\mathrm{CG}}\mathbf{v}^{(k)} + \rho_{k+1}\mathbf{r}^{(k+1)}$ *with* $\beta_{k+1}^{\mathrm{CG}} = m/(m + \alpha_k\delta_k)$ *and* $\rho_{k+1} = \alpha_k/(m + \alpha_k\delta_k)$*, which is unconditionally dissipative for* $\delta_k \geq 0$*. Here the coefficient* $\beta_{k+1}^{\mathrm{CG}}$ *is unrelated to the Newmark parameter* β *in* (7)*: the superscript "CG" is meant to prevent confusion. Thus* $\mathbf{v}^{(k+1)} \in \mathrm{span}\{\mathbf{r}^{(k+1)}, \mathbf{v}^{(k)}\}$ *as in the usual CG recurrence, but the coefficient in front of* $\mathbf{r}^{(k+1)}$ *is naturally* α_k/m*. Since* m *is constant, it can be absorbed into the time step, but keeping it explicit clarifies dimensions and reinforces the interpretation of* $\beta_{k+1}^{\mathrm{CG}}$ *as a discrete "memory" (damping) factor. The matrix-valued analogue of this viewpoint is precisely* $(3)_4$*, where* M *and* C *play the role of mass and damping operators.*

Remark 3.2 (From mass/damping matrices to a PCG-like structure)**.** *The scalar model in Remark 3.1 extends verbatim to the matrix-valued mass and damping operators appearing in* $(3)_4$*, namely*

$$\dot{\mathbf{u}} = \mathbf{v}, \qquad \mathbf{r} = \mathbf{b} - A\mathbf{u}, \qquad M\,\dot{\mathbf{v}} = \mathbf{r} - C\,\mathbf{v},$$

where M *maps accelerations to forces and* C *is a (typically SPD) damping operator. A key point is that the update of* $\mathbf{v}$ *always involves an* application of an inverse *of a matrix built from* M *and* C*, hence a natural preconditioning action on the residual.*

(Explicit/semi-implicit Euler). *With a step size* $\alpha_k = t_{k+1} - t_k$*, freezing* $\mathbf{v}$ *over the step in the first two equations gives*

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \alpha_k\,\mathbf{v}^{(k)}, \qquad \mathbf{r}^{(k+1)} = \mathbf{r}^{(k)} - \alpha_k\,A\,\mathbf{v}^{(k)}.$$

Discretizing $M\dot{\mathbf{v}} = \mathbf{r} - C\mathbf{v}$ by a forward Euler step and using the updated residual $\mathbf{r}^{(k+1)}$ yields

$$\mathbf{v}^{(k+1)} = \Big(I - \alpha_k M^{-1}C\Big)\,\mathbf{v}^{(k)} + \alpha_k\, M^{-1}\mathbf{r}^{(k+1)}. \tag{12}$$

Defining the preconditioned residual $\mathbf{z}^{(k+1)} := M^{-1}\mathbf{r}^{(k+1)}$, (12) *shows that*

$$\mathbf{v}^{(k+1)} \in \operatorname{span}\{\mathbf{z}^{(k+1)}, \mathbf{v}^{(k)}\},$$

which is the structural hallmark of a PCG-like two-term recurrence (compare with $\mathbf{v}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta_{k+1}\mathbf{v}^{(k)}$ in the standard preconditioned CG algorithm).

(Backward Euler on the damping law). *Using backward Euler for the damping equation gives*

$$\Big(M + \alpha_k C\Big)\,\mathbf{v}^{(k+1)} = M\,\mathbf{v}^{(k)} + \alpha_k\,\mathbf{r}^{(k+1)}, \tag{13}$$

hence

$$\mathbf{v}^{(k+1)} = \underbrace{(M + \alpha_k C)^{-1}M}_{=:T_k}\,\mathbf{v}^{(k)} + \alpha_k\,\underbrace{(M + \alpha_k C)^{-1}\mathbf{r}^{(k+1)}}_{=:\mathbf{z}^{(k+1)}}.$$

Again $\mathbf{v}^{(k+1)} \in \operatorname{span}\{\mathbf{z}^{(k+1)}, \mathbf{v}^{(k)}\}$, but now the residual is filtered by the effective (preconditioning) operator $(M + \alpha_k C)^{-1}$, which is unconditionally dissipative when $C \succeq 0$.

Connection with PCG. *For a fixed SPD preconditioner P, the (left-)preconditioned CG method for $A\mathbf{u} = \mathbf{b}$ can be written as*

$$\mathbf{z}^{(k)} = P^{-1}\mathbf{r}^{(k)}, \qquad \mathbf{v}^{(k+1)} = \mathbf{z}^{(k+1)} + \beta_{k+1}\mathbf{v}^{(k)}, \qquad \mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \alpha_k\mathbf{v}^{(k)},$$

with the classical choices

$$\alpha_k = \frac{(\mathbf{r}^{(k)})^T\mathbf{z}^{(k)}}{(\mathbf{v}^{(k)})^T A\mathbf{v}^{(k)}}, \qquad \beta_{k+1} = \frac{(\mathbf{r}^{(k+1)})^T\mathbf{z}^{(k+1)}}{(\mathbf{r}^{(k)})^T\mathbf{z}^{(k)}}.$$

The dynamical discretizations above provide the same two-term span *for the direction update with a natural preconditioned residual $\mathbf{z}^{(k)}$ (taking $P = M$ in (12), or $P_k = M + \alpha_k C$ in (13)). Recovering the* exact *PCG coefficients α_k, β_{k+1} amounts to enforcing the usual orthogonality/conjugacy conditions (a Galerkin projection in the A-inner product), which is a discrete constraint not built into the constant-coefficient time stepping.*

3.2 Error propagation and the role of M and C

Let $\mathbf{w}^{(k)} = \mathbf{u}^{(k)} - \overline{\mathbf{u}}$ be the algebraic error and let $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{u}^{(k)} = -A\mathbf{w}^{(k)}$. Equations (10)–(11) define a linear homogeneous recurrence for the pair $(\mathbf{w}^{(k)}, \mathbf{v}^{(k)})$, hence an iteration matrix whose spectral properties determine convergence. At this level, M plays the role of a *left preconditioner*: the stiffness matrix only appears through the product $M^{-1}A$ (or, more generally, through the pencil (A, M)), so choosing M such that $M^{-1}A$ has a favorable spectrum is expected to accelerate the iteration. The damping matrix C (together with γ and the time steps τ_k) controls how aggressively high-frequency components are attenuated.

3.3 Commuting damping and a mode-wise three-term recurrence

The previous discussion becomes particularly transparent when the damping operator *commutes* with the stiffness operator in the generalized eigenbasis of the pencil (A, M). Assume throughout that A and M are symmetric positive definite, so that there exists an M–orthonormal basis $\{\boldsymbol{\phi}_i\}_{i=1}^n$ and real eigenvalues $\lambda_i > 0$ such that

$$A\,\boldsymbol{\phi}_i = \lambda_i\, M\,\boldsymbol{\phi}_i, \qquad \boldsymbol{\phi}_i^T M\,\boldsymbol{\phi}_j = \delta_{ij}. \tag{14}$$

If, in addition, C is of *Rayleigh* form

$$C = a_0\, M + a_1\, A \qquad (a_0 \geq 0,\ a_1 \geq 0), \tag{15}$$

then $M^{-1}C$ is diagonal in the same basis and each generalized mode evolves independently.

Let $\mathbf{w}^{(k)} = \mathbf{u}^{(k)} - \overline{\mathbf{u}}$ denote the error and expand it as $\mathbf{w}^{(k)} = \sum_{i=1}^n \xi_i^{(k)}\,\boldsymbol{\phi}_i$ (with $\xi_i^{(k)} = \boldsymbol{\phi}_i^T M\,\mathbf{w}^{(k)}$). For $\beta = 0$, constant time step $\tau_k \equiv \tau$, and the choice $\gamma = \frac{1}{2}$, the homogeneous Newmark scheme applied to (3) reduces, for each mode, to the scalar recursion

$$\xi_i^{(k+2)} = p(\lambda_i)\,\xi_i^{(k+1)} + q(\lambda_i)\,\xi_i^{(k)}, \tag{16}$$

where the (mode-dependent) coefficients are

$$p(\lambda) = \frac{4 - 2\lambda\tau^2}{2 + \tau\, c(\lambda)}, \qquad q(\lambda) = \frac{\tau\, c(\lambda) - 2}{2 + \tau\, c(\lambda)}, \qquad c(\lambda) = a_0 + a_1\lambda. \tag{17}$$

Thus the iteration acts as a *spectral filter* on the modes of $M^{-1}A$: the choice of M shapes the eigenvalue interval $\{\lambda_i\}$, while (a_0, a_1) and τ control the amplification factors through (17). In the undamped case $a_0 = a_1 = 0$, (16) becomes the well-known central-difference recurrence [2, 3] $\xi_i^{(k+2)} = (2 - \lambda_i\tau^2)\xi_i^{(k+1)} - \xi_i^{(k)}$. When $a_0 > 0$ and $a_1 = 0$, the coefficients are independent of λ except for the term $\lambda\tau^2$ in p, and one recovers a standard momentum/heavy-ball behavior [4] (made explicit in Proposition 3.2).

Lemma 3.1 (Schur stability of a second-order recurrence)**.** *Consider the scalar linear recurrence*

$$x^{(k+2)} = p\,x^{(k+1)} + q\,x^{(k)}, \tag{18}$$

with real coefficients (p, q). Let $r_{1,2}$ be the roots of the characteristic polynomial $z^2 - pz - q = 0$. Then the following are equivalent:

(i) $|r_1| < 1$ *and* $|r_2| < 1$ *(Schur stability);*

(ii) $|q| < 1$ *and* $1 - q > |p|$.

Proof. This is the standard Jury (Schur) criterion for a real quadratic polynomial; see, e.g., [11]. Consider the monic polynomial

$$\pi(z) = z^2 + c_1 z + c_0,$$

with real coefficients (c_0, c_1). The Jury test states that both roots of π lie strictly inside the unit disk if and only if

$$|c_0| < 1, \qquad 1 + c_1 + c_0 > 0, \qquad 1 - c_1 + c_0 > 0.$$

Applying this to $z^2 - pz - q = 0$ (i.e., $c_1 = -p$ and $c_0 = -q$) yields $|q| < 1$, and

$$1 - p - q > 0, \qquad 1 + p - q > 0,$$

which are equivalent to $1 - q > |p|$. □

Proposition 3.1 (Step-size condition and convergence to the static solution)**.** *Assume A and M are symmetric positive definite and C is of Rayleigh form* (15) *with $(a_0, a_1) \neq (0, 0)$. Let $\{\lambda_i\}$ be the generalized eigenvalues* (14) *and set $\lambda_{\max} = \max_i \lambda_i = \lambda_{\max}(M^{-1}A)$. For the Newmark choice $\beta = 0$, $\gamma = \frac{1}{2}$ and a constant time step $\tau > 0$ satisfying*

$$0 < \tau^2 \lambda_{\max} < 4, \tag{19}$$

the mode-wise recurrence (16) *is Schur stable for every i, hence $\xi_i^{(k)} \to 0$ as $k \to \infty$ and therefore $\mathbf{u}^{(k)} \to \overline{\mathbf{u}}$ for any initial data.*

Proof. Fix a mode i and abbreviate $\lambda = \lambda_i$ and $d = \tau\, c(\lambda) = \tau(a_0 + a_1\lambda)$. Since $(a_0, a_1) \neq (0, 0)$ and $\lambda > 0$, we have $d > 0$. Using (17),

$$q(\lambda) = \frac{d - 2}{d + 2}, \qquad p(\lambda) = \frac{4 - 2\lambda\tau^2}{d + 2}.$$

Then $|q(\lambda)| < 1$ holds for all $d > 0$. Moreover,

$$1 - q(\lambda) = \frac{4}{d + 2}, \qquad |p(\lambda)| = \frac{|4 - 2\lambda\tau^2|}{d + 2}.$$

By Lemma 3.1, Schur stability is therefore equivalent to $4 > |4 - 2\lambda\tau^2|$, i.e. to $0 < \lambda\tau^2 < 4$. Condition (19) implies $0 < \lambda_i\tau^2 < 4$ for every i, hence all modes decay and $\mathbf{u}^{(k)} \to \overline{\mathbf{u}}$. □

Theorem 3.1 (Rayleigh damping: modal recursion and contraction)**.** *Assume $A \succ 0$ and $M \succ 0$, and let $C = a_0 M + a_1 A$ with $a_0 \geq 0$, $a_1 \geq 0$ and $(a_0, a_1) \neq (0, 0)$. Consider the explicit Newmark scheme with $\beta = 0$, $\gamma = \frac{1}{2}$ and constant time step $\tau > 0$. Let $\{\lambda_i\}_{i=1}^n$ be the generalized eigenvalues of (A, M) and set $L = \lambda_{\max}(M^{-1}A) = \max_i \lambda_i$. Then each modal amplitude $\xi_i^{(k)}$ of the error satisfies the two-step recursion* (16) *with coefficients* (17)*. If*

$$0 < \tau^2 L < 4,$$

then the characteristic roots of $z^2 - p(\lambda_i)z - q(\lambda_i) = 0$ lie in the open unit disk for every i and hence $\mathbf{u}^{(k)} \to \overline{\mathbf{u}}$ as $k \to \infty$ for any initial data. Moreover, in the underdamped regime $p(\lambda)^2 + 4q(\lambda) < 0$ the roots are complex conjugates and their modulus is

$$|z_\pm(\lambda)| = \sqrt{-q(\lambda)} = \sqrt{\frac{2 - \tau\, c(\lambda)}{2 + \tau\, c(\lambda)}}, \qquad c(\lambda) = a_0 + a_1\lambda,$$

so that the damping controls the asymptotic decay mode-wise, while the step-size restriction is dictated by the interval of $M^{-1}A$.

3.3.1 Mode-wise spectral radius and parameter tuning

For each mode i, define the dimensionless quantities

$$x_i := \tau^2\, \lambda_i, \qquad d_i := \tau\, c(\lambda_i) = \tau\,(a_0 + a_1\lambda_i). \tag{20}$$

In terms of (x_i, d_i), the scalar recurrence (16) reads

$$p(\lambda_i) = \frac{2(2 - x_i)}{2 + d_i}, \qquad q(\lambda_i) = \frac{d_i - 2}{2 + d_i}, \tag{21}$$

and the characteristic roots are

$$r_\pm(\lambda_i) = \frac{p(\lambda_i) \pm \sqrt{p(\lambda_i)^2 + 4q(\lambda_i)}}{2}. \tag{22}$$

A short computation shows that the discriminant simplifies to

$$p(\lambda)^2 + 4q(\lambda) = \frac{4}{(2 + d(\lambda))^2}\left(x(\lambda)^2 - 4x(\lambda) + d(\lambda)^2\right), \tag{23}$$

where $x(\lambda) = \tau^2\lambda$ and $d(\lambda) = \tau(a_0 + a_1\lambda)$.

Underdamped regime (complex roots). The roots are complex conjugate if and only if

$$(x - 2)^2 < 4 - d^2 \qquad \text{(necessarily } 0 \le d < 2\text{)}. \tag{24}$$

In this case, the modulus is determined by the product $r_+ r_- = -q$ and one obtains the explicit per-step amplification factor

$$|r_\pm(\lambda)| = \sqrt{-q(\lambda)} = \sqrt{\frac{2 - d(\lambda)}{2 + d(\lambda)}}. \tag{25}$$

In particular, in the proportional damping case $C = \eta M$ (so that $d(\lambda) \equiv \eta\tau$), the damping factor (25) is *mode-independent* whenever (24) holds.

Overdamped regime (real roots). If (24) fails, the roots are real and the spectral radius for that mode is

$$\rho(\lambda) := \max\{|r_+(\lambda)|, |r_-(\lambda)|\} = \frac{|p(\lambda)| + \sqrt{p(\lambda)^2 + 4q(\lambda)}}{2}. \tag{26}$$

The step-size condition (19) guarantees $\rho(\lambda) < 1$ for all $\lambda \in (0, \lambda_{\max}]$, but the actual convergence rate is governed by the worst-case value $\max_i \rho(\lambda_i)$.

Tuning guidelines. Equation (25) shows that increasing $d(\lambda)$ (i.e. increasing a_0 and/or a_1) decreases the modulus in the underdamped region, thus suppressing oscillations; however, too large damping may push modes into the real-root regime and slow down asymptotic decay. The Rayleigh term $a_1 A$ makes $d(\lambda)$ grow with λ, providing stronger damping of high-frequency components (large λ) while leaving low-frequency components less damped.

Remark 3.3 (Matching classical heavy-ball parameters in the proportional case)**.** *In the proportional case $C = \eta M$,* (10)–(11) *induces a preconditioned heavy-ball iteration with parameters*

$$\alpha = \frac{2\tau^2}{2+d}, \qquad \mu = \frac{2-d}{2+d}, \qquad d = \eta\tau. \tag{27}$$

If one has estimates $0 < m \le \lambda_i \le L$ for the spectrum of $M^{-1}A$, one may match the classical "optimal" heavy-ball parameters on $[m, L]$,

$$\alpha^\star = \frac{4}{\left(\sqrt{L} + \sqrt{m}\right)^2}, \qquad \mu^\star = \left(\frac{\sqrt{L} - \sqrt{m}}{\sqrt{L} + \sqrt{m}}\right)^2, \tag{28}$$

by enforcing $\mu = \mu^\star$ and $\alpha = \alpha^\star$ through (27)*, that is*

$$d = \frac{2(1-\mu^\star)}{1+\mu^\star}, \qquad \tau^2 = \frac{\alpha^\star}{2}(2+d), \qquad \eta = \frac{d}{\tau}. \tag{29}$$

This provides a concrete "preconditioned design" recipe once rough spectral bounds for $M^{-1}A$ are available.

Remark 3.4 (Preconditioning viewpoint)**.** *The step-size constraint* (19) *depends only on the spectral radius of the preconditioned operator $M^{-1}A$. Choosing M as an efficient SPD preconditioner for A (so that the spectrum of $M^{-1}A$ is clustered) enlarges the admissible time step τ and typically improves the convergence rate of the associated iteration.*

Remark 3.5 (Beyond Rayleigh damping)**.** *The decoupling argument applies more generally whenever $M^{-1}C$ is a (low-degree) polynomial in $M^{-1}A$, or, equivalently, whenever C is in the algebra generated by M and A. In such cases one obtains a scalar recursion of the form* (16) *with coefficients that are rational functions of λ, and the design of (M, C, τ) can be guided by the desired mode-wise amplification on a target eigenvalue interval.*

3.4 A clean case: proportional damping yields a preconditioned heavy-ball method

A particularly transparent situation is obtained by assuming

$$C = \eta\, M \qquad (\eta > 0), \tag{30}$$

a constant time step $\tau_k \equiv \tau$, and the choice $\gamma = \frac{1}{2}$ (often referred to as the "central difference" value in structural dynamics). In this case the iteration can be written solely in terms of $\mathbf{u}$ (or $\mathbf{w}$) and becomes a standard *momentum* method for (1).

Proposition 3.2 (Equivalence with a preconditioned heavy-ball iteration)**.** *Assume* (30)*,* $\tau_k \equiv \tau$*,* $\gamma = \frac{1}{2}$*, and* $\beta = 0$*. Then the displacement iterates* $\{\mathbf{u}^{(k)}\}$ *produced by Newmark satisfy the two-step recurrence*

$$\mathbf{u}^{(k+2)} = \mathbf{u}^{(k+1)} + \mu\,\big(\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}\big) + \alpha\, M^{-1}\big(\mathbf{b} - A\,\mathbf{u}^{(k+1)}\big), \tag{31}$$

with parameters

$$\alpha = \frac{2\,\tau^2}{2 + \eta\,\tau}, \qquad \mu = \frac{2 - \eta\,\tau}{2 + \eta\,\tau}. \tag{32}$$

Equivalently, in terms of the error $\mathbf{w}^{(k)} = \mathbf{u}^{(k)} - \overline{\mathbf{u}}$*,*

$$\mathbf{w}^{(k+2)} = \big(I - \alpha\, M^{-1}A\big)\,\mathbf{w}^{(k+1)} + \mu\,\big(\mathbf{w}^{(k+1)} - \mathbf{w}^{(k)}\big). \tag{33}$$

Proof. The proof is easiest in modal coordinates. Since $C = \eta M$, the matrices $M^{-1}A$ and $M^{-1}C = \eta I$ commute and share the eigenvectors of the pencil (A, M) in Appendix A. Projecting the homogeneous scheme (i.e., $\mathbf{b} = 0$ and $\mathbf{u} \equiv \mathbf{w}$) onto any generalized eigenvector of (A, M) with eigenvalue $\lambda = \omega^2$ reduces (10)–(11) (with $\gamma = \frac{1}{2}$ and $\tau_k \equiv \tau$) to a 2×2 linear recurrence for the scalar pair $(y^{(k)}, \dot{y}^{(k)})$. Eliminating $\dot{y}^{(k)}$ yields

$$y^{(k+2)} = (1 + \mu - \alpha\,\lambda)\, y^{(k+1)} - \mu\, y^{(k)},$$

with α, μ given by (32). Returning to vector form gives (33), and adding back the affine shift $\overline{\mathbf{u}}$ yields (31). □

Remark 3.6 (Preconditioning viewpoint)**.** *In* (31)*–*(33)*, the matrix* M *plays the role of a (left) preconditioner. If* $M \approx A$ *is an approximation for which applying* M^{-1} *(i.e., solving with* M*) is inexpensive, then the spectrum of* $M^{-1}A$ *is better clustered and the heavy-ball iteration* (31) *can converge substantially faster.*

Remark 3.7 (Parameter tuning and relation to CG)**.** *For SPD systems, if one has bounds* $\lambda_{\min} \le \lambda(M^{-1}A) \le \lambda_{\max}$*, then classical momentum methods suggest choosing* (α, μ) *as functions of* $(\lambda_{\min}, \lambda_{\max})$ *to obtain near-optimal polynomial damping over the whole spectrum. This yields a Chebyshev-type acceleration [5, 8]. Conjugate gradients can be viewed as producing, in an adaptive way, an optimal polynomial in the* A *[6, 7, 8]-norm without requiring explicit spectral bounds; connecting* (31) *to CG therefore typically requires allowing step-dependent parameters (or additional orthogonality conditions), which we do not pursue here.*

Theorem 3.2 (Optimal contraction for proportional damping on a spectral interval)**.** *Assume A and M are symmetric positive definite and that the generalized eigenvalues of (A, M) satisfy $0 < m \le \lambda_i \le L$. Consider the explicit Newmark iteration with $\beta = 0$, $\gamma = \frac{1}{2}$, constant step τ, and proportional damping $C = \eta M$. Choose (τ, η) through the mapping* (36) *induced by the classical "optimal" heavy-ball parameters* (35) *on $[m, L]$. Then the error satisfies*

$$\|\mathbf{w}^{(k)}\|_M \le \rho_\star^k \left(\|\mathbf{w}^{(0)}\|_M + \|\mathbf{w}^{(1)}\|_M \right) \qquad \textit{with} \qquad \rho_\star = \frac{\sqrt{L} - \sqrt{m}}{\sqrt{L} + \sqrt{m}} \in (0, 1),$$

where $\|x\|_M^2 = x^T M x$. Equivalently, the heavy-ball form (33) *has a mode-wise spectral radius bounded by $\rho_\star$ on $[m, L]$.*

Proof. Project (33) onto an M–orthonormal eigenvector of $M^{-1}A$ with eigenvalue $\lambda \in [m, L]$. The scalar error component $\xi^{(k)}$ satisfies

$$\xi^{(k+2)} = (1 + \mu - \alpha\lambda)\, \xi^{(k+1)} - \mu\, \xi^{(k)},$$

with $(\alpha, \mu) = (\alpha^\star, \mu^\star)$ from (35). For these parameters, the characteristic roots are complex conjugates for all $\lambda \in [m, L]$ and their modulus equals $\sqrt{\mu^\star} = (\sqrt{L} - \sqrt{m})/(\sqrt{L} + \sqrt{m}) = \rho_\star$. Therefore each mode decays by at most $\rho_\star$ per step in the M–norm, and summing over modes yields the stated bound. □

Corollary 3.1 (Iteration complexity)**.** *Under the assumptions of Theorem 3.2, the number of Newmark steps needed to reduce the M–norm of the error by a factor $\varepsilon \in (0, 1)$ is $k = \mathcal{O}(\sqrt{\kappa}\, \log(1/\varepsilon))$, where $\kappa = L/m$ is the condition number of $M^{-1}A$.*

3.5 Beyond proportional damping

The proportional model (30) is primarily meant to expose the algebraic structure. More general choices of C (e.g., Rayleigh damping $C = a_0 M + a_1 A$ or other commuting constructions) lead to similar modal recurrences in which each mode is damped by a polynomial in the generalized eigenvalue ω_i^2. This is precisely the mechanism that makes the "dynamic" viewpoint useful for designing iterative solvers: by selecting M and C one shapes the iteration polynomial applied to the preconditioned operator $M^{-1}A$.

3.6 A simple design workflow

The previous results show that the convergence of the Newmark time-marching iteration is controlled by the spectrum of the (preconditioned) operator $M^{-1}A$ together with the dimensionless parameters $x = \tau^2 \lambda$ and $d(\lambda) = \tau\, c(\lambda)$, where $c(\lambda)$ is the modal damping induced by C. This suggests a practical "design" viewpoint: choose M (as a preconditioner), estimate a spectral interval for $M^{-1}A$, and then tune (τ, C) to obtain a desirable spectral radius.

Step 1: choose M as a preconditioner. Let M be any SPD matrix such that applying M^{-1} (i.e., solving linear systems with M) is inexpensive (exactly or approximately). Typical choices include diagonal scaling, incomplete factorizations, block preconditioners, or multilevel approximations. [10, 8, 9] All results are naturally expressed in terms of the generalized eigenvalues λ_i of (A, M), i.e. of $M^{-1}A$.

Step 2: estimate spectral bounds. Let $0 < m \le \lambda_{\min}(M^{-1}A)$ and $L \ge \lambda_{\max}(M^{-1}A)$ be coarse bounds. In practice, a few steps of power iteration (for L) or Lanczos (for m and L) are often sufficient; only matrix–vector products with A and solves with M are needed.

Step 3: pick a stable time step. For the explicit Newmark case $\beta = 0$ (with $\gamma = \frac{1}{2}$) the basic stability requirement reads

$$0 < \tau^2 L < 4, \tag{34}$$

so that a safe choice is $\tau = 2\,\theta/\sqrt{L}$ with $\theta \in (0, 1)$.

Step 4a: proportional damping $C = \eta M$ (heavy-ball form). If $C = \eta M$, the iteration on $\mathbf{u}$ can be written as a preconditioned heavy-ball scheme (see §3.4):

$$\mathbf{u}^{(k+2)} = \mathbf{u}^{(k+1)} + \mu\big(\mathbf{u}^{(k+1)} - \mathbf{u}^{(k)}\big) + \alpha\, M^{-1}(\mathbf{b} - A\mathbf{u}^{(k+1)}).$$

Given a spectral interval $[m, L]$, one may use the classical "optimal" heavy-ball parameters for quadratic problems,

$$\alpha^\star = \frac{4}{\big(\sqrt{L} + \sqrt{m}\big)^2}, \qquad \mu^\star = \left(\frac{\sqrt{L} - \sqrt{m}}{\sqrt{L} + \sqrt{m}}\right)^2. \tag{35}$$

These map to Newmark parameters through

$$d^\star := \eta\tau = \frac{2(1 - \mu^\star)}{1 + \mu^\star}, \qquad \tau = \sqrt{\frac{2\alpha^\star}{1 + \mu^\star}}, \qquad \eta = \frac{d^\star}{\tau}, \tag{36}$$

which automatically satisfies (34) since $\tau^2 = 4/(L + m)$ in this case. The resulting worst-case contraction factor on $[m, L]$ is given by Theorem 3.2.

Step 4b: Rayleigh damping $C = a_0 M + a_1 A$ (mode-selective damping). With Rayleigh damping, $d(\lambda) = \tau(a_0 + a_1\lambda)$ increases linearly with λ. This allows one to damp high-frequency modes more strongly while keeping low modes mildly damped. A simple two-point design is obtained by prescribing

$$d_m := d(m), \qquad d_L := d(L),$$

and solving for (a_0, a_1):

$$a_1 = \frac{d_L - d_m}{\tau(L-m)}, \qquad a_0 = \frac{d_m}{\tau} - a_1\, m. \tag{37}$$

Typical choices keep d_m small (to avoid overdamping low modes) and take d_L closer to 2 to quickly damp high modes, while maintaining (34). Setting $a_1 = 0$ (i.e. $d_L = d_m$) reduces to proportional damping.

Corollary 3.2 (Rayleigh damping as a smoother for high-frequency modes)**.** *Assume* (34) *and consider Rayleigh damping with parameters chosen so that* $0 \le d(\lambda) = \tau(a_0 + a_1\lambda) < 2$ *for* $\lambda \in [m, L]$*. For any subset of modes such that* (24) *holds (in particular, for sufficiently large* λ *if* $a_1 > 0$*), the corresponding error components satisfy the explicit bound*

$$|\xi^{(k)}(\lambda)| \le \left(\sqrt{\frac{2 - d(\lambda)}{2 + d(\lambda)}}\right)^k \Big(|\xi^{(0)}(\lambda)| + |\xi^{(1)}(\lambda)|\Big).$$

Thus, selecting $a_1 > 0$ *to make* $d(\lambda)$ *increase with* λ *yields stronger attenuation of high-frequency components, a behavior analogous to smoothing in multilevel methods.*

Remark 3.8 (Varying parameters and Chebyshev-like acceleration)**.** *When* $C = 0$ *and* $\beta = 0$*, stability imposes only* $\tau_k^2\lambda \le 4$ *mode-wise. By letting* τ_k *(or, equivalently, the residual step length) vary with* k*, one can realize polynomial filters closely related to Chebyshev semi-iteration on the interval* $[m, L]$*. This viewpoint also clarifies the connection with Krylov methods: CG adaptively builds near-optimal polynomials on* $[m, L]$*, whereas the Newmark-based iteration produces a prescribed polynomial through* (τ_k, C)*.*

4 A compact iteration viewpoint

For the reader's convenience, we summarize here the algebraic form of the explicit Newmark scheme ($\beta = 0$) as a linear iteration for the error in (1). This provides a single "entry point" for analysis (via iteration matrices or mode-wise recurrences) and for algorithmic design (via the choice of (M, C, τ_k)).

4.1 State-space iteration matrix for the homogeneous error dynamics

Consider the homogeneous error variables $\mathbf{w}^{(k)} = \mathbf{u}^{(k)} - \overline{\mathbf{u}}$ and $\mathbf{v}^{(k)} \approx \dot{\mathbf{u}}(t_k)$, so that $\mathbf{r}^{(k)} = \mathbf{b} - A\mathbf{u}^{(k)} = -A\mathbf{w}^{(k)}$. For a fixed time step $\tau_k \equiv \tau$ and fixed $\gamma \in [0, 1]$, the explicit scheme (10)–(11) defines a linear recurrence

$$\begin{pmatrix} \mathbf{w}^{(k+1)} \\ \mathbf{v}^{(k+1)} \end{pmatrix} = \mathcal{G} \begin{pmatrix} \mathbf{w}^{(k)} \\ \mathbf{v}^{(k)} \end{pmatrix}, \tag{38}$$

with a $2n \times 2n$ iteration matrix $\mathcal{G}$ whose blocks depend only on M, C, A, τ, γ. To make this explicit, rewrite (10) (using $\mathbf{r}^{(k)} = -A\mathbf{w}^{(k)}$) as

$$\mathbf{w}^{(k+1)} = \underbrace{\left(I - \tfrac{\tau^2}{2} M^{-1} A\right)}_{=:G_{11}} \mathbf{w}^{(k)} + \underbrace{\left(\tau I - \tfrac{\tau^2}{2} M^{-1} C\right)}_{=:G_{12}} \mathbf{v}^{(k)}. \tag{39}$$

Next, insert $\mathbf{r}^{(k)} = -A\mathbf{w}^{(k)}$ and $\mathbf{r}^{(k+1)} = -A\mathbf{w}^{(k+1)}$ into (11) and solve for $\mathbf{v}^{(k+1)}$:

$$\mathbf{v}^{(k+1)} = (M + \gamma\tau C)^{-1} \Big[(M - (1-\gamma)\tau C)\mathbf{v}^{(k)} - \tau \big(\gamma A \mathbf{w}^{(k+1)} + (1-\gamma) A \mathbf{w}^{(k)} \big) \Big]. \tag{40}$$

Combining (39)–(40) gives (38) with

$$\mathcal{G} = \begin{pmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{pmatrix}, \qquad \begin{cases} G_{21} = -(M + \gamma\tau C)^{-1} \tau \left(\gamma A G_{11} + (1-\gamma) A \right), \\ G_{22} = (M + \gamma\tau C)^{-1} \Big[(M - (1-\gamma)\tau C) - \tau\, \gamma A G_{12} \Big]. \end{cases} \tag{41}$$

When $\mathcal{G}$ is constant (fixed τ), linear convergence is governed by the spectral radius $\rho(\mathcal{G})$. The Rayleigh and proportional settings singled out in §3.3–§3.4 correspond precisely to cases where (38) decouples mode-wise and yields explicit scalar recurrences. In the variable-step setting τ_k, (10)–(11) generate a product of matrices $\mathcal{G}_k$; this is the natural framework for interpreting the method as a polynomial (or rational) filter on the spectrum of $M^{-1}A$.

5 Scope, assumptions, and limitations

The "dynamics-to-iteration" correspondence is cleanest under the standing SPD assumptions used throughout. For robustness (and to guide future extensions), it is useful to keep the following points explicit.

SPD structure. Most statements in the main text assume $A \succ 0$ and $M \succ 0$. The equilibrium interpretation is still meaningful for symmetric indefinite problems, but stability and convergence then depend on how the negative spectrum is handled (e.g., shifted dynamics or stabilized saddle-point formulations), and CG/PCG no longer apply directly.

Damping and asymptotic decay. For the continuous system (5), $C \succ 0$ yields global asymptotic stability (Appendix B). If $C \succeq 0$ one generally obtains Lyapunov stability; asymptotic decay requires excluding undamped modes. At the discrete level, using backward Euler on the damping law (Remarks 3.1–3.2) is unconditionally dissipative in the $C \succeq 0$ case.

Commutativity assumptions. The sharp mode-wise analysis in §3.3 relies on $M^{-1}C$ being diagonalizable in the generalized eigenbasis of (A, M) (e.g., Rayleigh/Caughey damping). Without this property one can still analyze (38) via matrix norms or by bounding non-normality effects, but closed-form scalar recurrences are typically unavailable.

What is (and is not) "PCG" here. The Newmark-induced iterations naturally generate a two-term (CG/PCG-like) *span* for the direction update (Remark 3.2), with a built-in preconditioning action given by M^{-1} or $(M+\tau C)^{-1}$. Recovering the *exact* PCG coefficients is a separate step: it amounts to enforcing the discrete Galerkin orthogonality/conjugacy conditions that characterize Krylov methods.

6 Numerical experiments

This section illustrates the practical meaning of the "dynamics-to-iteration" viewpoint through a small set of *reproducible* SPD test problems. The aim is *not* to compete with optimized Krylov solvers, but to document, in a controlled way, how the Newmark-induced iteration behaves as a stationary filter whose design parameters are (M, C, τ).

6.1 Test problems and discretizations

All tests are posed on $\Omega = (0,1)^2$ with homogeneous Dirichlet boundary conditions. We discretize on a uniform $n \times n$ grid of *interior* points ($n = 80$ in all runs reported below), so that the linear system dimension is $N = n^2$. To remove PDE discretization error from the comparison we use a manufactured solution strategy: we build a smooth "true" vector $\mathbf{u}_\star$ as a small combination of sine modes and set $\mathbf{b} = A\mathbf{u}_\star$. All methods start from $\mathbf{u}^{(0)} = \mathbf{0}$ (hence $\mathbf{r}^{(0)} = \mathbf{b}$) and stop when $\|\mathbf{r}^{(k)}\|_2/\|\mathbf{b}\|_2 \le 10^{-10}$ or when $k = \texttt{maxit} = 5000$.

(T1) Poisson (5-point FD). A is the standard 5-point finite-difference Laplacian.

(T2) Heterogeneous diffusion (checkerboard). A is a finite-volume discretization of $-\nabla\cdot(\kappa\nabla u)$ with a 8×8 checkerboard coefficient taking values $\kappa_{\mathrm{lo}} = 1$ and $\kappa_{\mathrm{hi}} = 10^4$.

(T3) Anisotropic diffusion (rotated tensor). A corresponds to a constant anisotropic tensor $\mathbf{K} = R(\theta)\mathrm{diag}(1,\varepsilon)R(\theta)^T$ with $(\theta,\varepsilon) = (\pi/6, 10^{-3})$, discretized by a standard 9-point FD stencil.

6.2 Compared methods

We compare (i) reference Krylov methods and (ii) Newmark-induced iterations.

- **CG/PCG (MATLAB `pcg`).** We use either no preconditioner (CG) or a symmetric SPD preconditioner P (PCG).
- **Newmark–prop.** The Newmark-induced stationary iteration with proportional damping $C = \eta M$ and $\beta = 0, \gamma = \frac{1}{2}$ (the explicit-displacement scheme analyzed in the paper).

- **Newmark–Rayleigh.** The same scheme but with Rayleigh damping $C = a_0 M + a_1 A$. For M non-diagonal, the velocity update is solved *inexactly* by an inner PCG (see below).

6.3 Preconditioners, parameter selection, and inner solves

We test three choices of "mass/preconditioner" M:

$$M = I, \qquad M = D := \mathrm{diag}(A) \ \text{(Jacobi)}, \qquad M = P_{\mathrm{IC}} := L_{\mathrm{IC}} L_{\mathrm{IC}}^T \ \text{(ICscaled)}.$$

Here P_{IC} is built by an incomplete Cholesky factorization of $D^{-1/2} A D^{-1/2}$ (MATLAB `ichol`, `ict`, with `droptol`$= 10^{-3}$, `diagcomp`$= 0.05$, and `michol`=`on`) followed by rescaling back to the original variables; this is the same preconditioner used for the PCG–ICscaled baseline.

Newmark–prop. We set $\eta\tau = 1$ (equivalently, $\eta = 1/\tau$) and choose τ from the stability bound $\tau^2 \lambda_{\max}(M^{-1}A) < 4$ by setting

$$\tau = \frac{2\theta}{\sqrt{\lambda_{\max}(M^{-1}A)}}, \qquad \theta = 0.95,$$

where $\lambda_{\max}(M^{-1}A)$ is estimated by a few `eigs` iterations when M is diagonal and by a power iteration on $P_{\mathrm{IC}}^{-1} A$ when $M = P_{\mathrm{IC}}$.

Newmark–Rayleigh. We set a *target damping ratio* $\zeta = 0.7$ and determine (a_0, a_1) from approximate spectral bounds $0 < m \le \sigma(M^{-1/2} A M^{-1/2}) \le L$ using the standard Rayleigh matching

$$a_0 = 2\zeta \, \frac{\sqrt{mL}}{\sqrt{m} + \sqrt{L}}, \qquad a_1 = 2\zeta \, \frac{1}{\sqrt{m} + \sqrt{L}}. \tag{42}$$

This yields moderate damping across the full spectrum while preserving the $M^{-1}A$-modal structure. We keep the same τ as in Newmark–prop.

Inner PCG for the velocity update (Rayleigh, $M = P_{\mathrm{IC}}$). With non-diagonal M, the velocity update requires solving

$$Q\, \mathbf{v}^{(k+1)} = \mathrm{rhs}, \qquad Q = (1 + \gamma\tau a_0) M + (\gamma\tau a_1) A,$$

with a variable tolerance `tol_in` scaled to the current outer residual. Specifically, at outer step k we set

$$\texttt{tol\_in}^{(k)} = \min(10^{-2}, \max(10^{-4},\, 0.1\, \|r_k\|_2 / \|b\|_2))$$

and solve until the relative inner residual drops below $\texttt{tol\_in}^{(k)}$. We precondition the inner solve by M (one application of M^{-1} per inner step), warm-start from $\mathbf{v}^{(k)}$, cap the inner iterations to 60, and track the total number of inner iterations.

	T1		T2		T3	
Method	it	final	it	final	it	final
CG	13	8.5×10^{-11}	2287	7.8×10^{-11}	569	9.7×10^{-11}
PCG–Jacobi	13	8.5×10^{-11}	555	9.2×10^{-11}	569	9.7×10^{-11}
PCG–ICscaled	46	5.5×10^{-11}	119	4.2×10^{-11}	91	8.0×10^{-11}
Newmark–prop ($M=I$)	5000	2.1×10^{-4}	5000	7.9×10^{-4}	5000	6.0×10^{-4}
Newmark–prop ($M=D$)	5000	2.1×10^{-4}	5000	1.4×10^{-4}	5000	6.0×10^{-4}
Newmark–prop ($M=P_{\mathrm{IC}}$)	367	9.4×10^{-11}	5000	6.9×10^{-5}	464	9.9×10^{-11}
Newmark–Rayleigh ($M=D$)	788	1.0×10^{-10}	5000	1.2×10^{-6}	896	9.8×10^{-11}
Newmark–Rayleigh ($M=P_{\mathrm{IC}}$)	124	7.9×10^{-11}	2626	9.9×10^{-11}	140	7.6×10^{-11}

Table 1: Iteration counts (outer iterations) and final relative residuals. For Newmark–Rayleigh with $M=P_{\mathrm{IC}}$ the total number of inner PCG steps is: T1: 215, T2: 2742, T3: 320.

6.4 Work units

As a hardware-agnostic proxy for cost, we count A–matvecs. Table 1 reports outer iteration counts and, when applicable, the total number of inner PCG steps; from these one can reconstruct A–matvec counts via the rules below. For PCG, one iteration corresponds to one A–matvec. For Newmark–prop, one outer step uses one A–matvec. For Newmark–Rayleigh, one outer step uses two A–matvecs (one for $A\mathbf{v}^{(k)}$, one for $A\mathbf{u}^{(k+1)}$), plus one A–matvec per inner PCG step.

6.5 Results

Figures 1–3 show residual histories. Table 1 summarizes iteration counts and (when applicable) the total number of inner PCG steps.

Discussion. Two qualitative behaviors are consistent with the theory developed above. First, proportional damping ($C=\eta M$) behaves as a *smoother*: it contracts high-frequency components but may stall before reaching a tight tolerance unless the preconditioner is strong enough (compare $M=I, D$ vs. $M=P_{\mathrm{IC}}$). Second, Rayleigh damping can be tuned to damp both ends of the spectrum and, when combined with a strong M, yields a robust stationary iteration that reaches the PCG tolerance even on the high-contrast test (T2), at the price of substantially more outer steps. In all reported runs, the inner PCG cost for the velocity update remained modest (of the same order as the number of outer iterations), suggesting that the inexact Rayleigh variant is feasible as a "preconditioner-in-the-loop" mechanism.

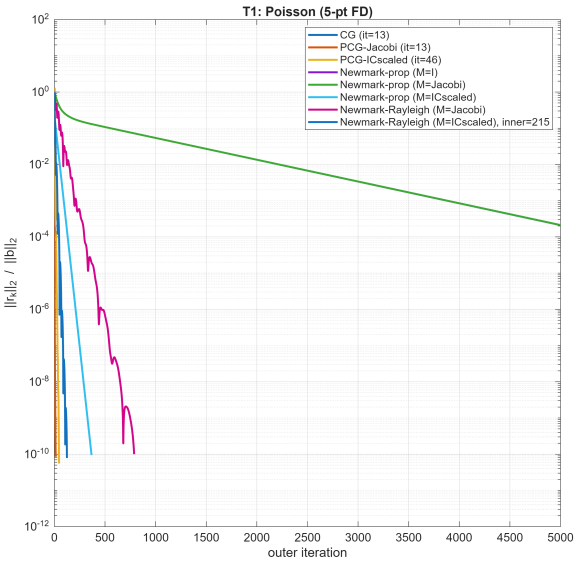

Figure 1: (T1) Poisson. Residual histories for CG/PCG and Newmark-induced iterations.

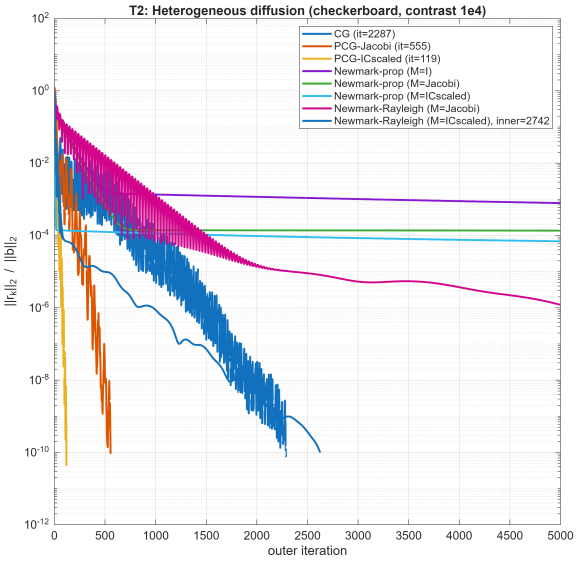

Figure 2: (T2) Heterogeneous diffusion (contrast 10^4). The proportional-damping Newmark iteration stalls at a relatively large plateau, while Rayleigh damping combined with $M = P_{\mathrm{IC}}$ reaches the Krylov tolerance (oscillations correspond to underdamped modes).

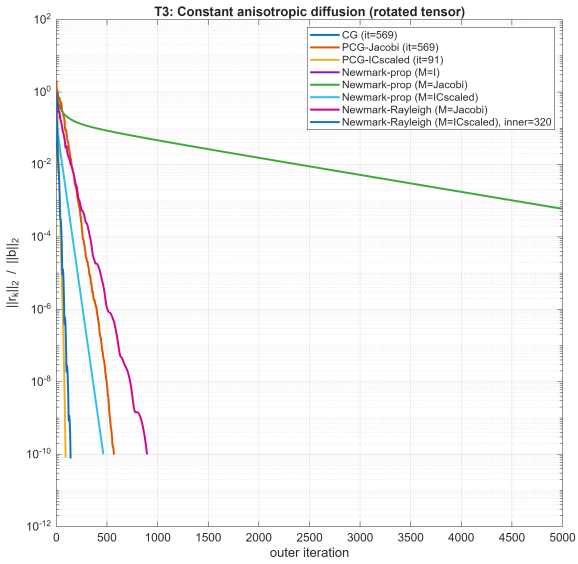

Figure 3: (T3) Anisotropic diffusion. Newmark–Rayleigh with $M = P_{\mathrm{IC}}$ converges robustly with a small total inner cost.

6.6 Sensitivity and scaling

To assess robustness beyond the default parameter set used in Table 1, we report two additional checks: (i) a damping-ratio sweep for the high-contrast test (T2), and (ii) a mild grid-refinement study.

Sensitivity to the target damping ratio ζ (T2). We keep the same preconditioner $M = P_{\mathrm{IC}}$ and the same stability-driven step size $\tau = 2\theta/\sqrt{\lambda_{\max}(M^{-1}A)}$ with $\theta = 0.95$, and vary only the Rayleigh coefficients (a_0, a_1) through the target damping ratio ζ in (42). Table 2 summarizes the outer iteration count, the final relative residual, and the total number of inner PCG steps (for the v-update) for a representative set of ζ values. In our experience, too small ζ

ζ	outer it	final relres	inner it (total)	notes
0.30	5000	2.07×10^{-9}	5167	maxit
0.50	3829	9.95×10^{-11}	3958	
0.70	2629	9.87×10^{-11}	2749	
0.90	2111	9.88×10^{-11}	2376	
1.20	2177	9.95×10^{-11}	3014	

Table 2: Damping-ratio sweep for T2 with Newmark–Rayleigh and $M = P_{\mathrm{IC}}$.

underdamps low-frequency modes and slows the tail convergence, while too large ζ overdamps and reduces the effective contraction per step. A moderate range (e.g. $\zeta \in [0.5, 1.0]$) typically balances these effects. In Table 2, $\zeta = 0.9$ yields the smallest outer iteration count among the tested values, while $\zeta = 0.3$ does not reach the target tolerance within the prescribed cap (`maxit`=5000).

Scaling with grid refinement (T2). We repeat the T2 experiment for increasing grid sizes n (hence $N = n^2$), using the same setup (manufactured $\mathbf{u}_\star$, $M = P_{\mathrm{IC}}$, and Rayleigh damping with $\zeta = 0.7$). Table 3 reports the iteration counts of PCG–ICscaled and Newmark–Rayleigh with $M = P_{\mathrm{IC}}$, together with the cumulative inner PCG steps for the latter. These results

n	N	PCG–ICscaled it	Rayleigh outer it	Rayleigh inner it (total)
40	1600	67	1697	1772
80	6400	119	2629	2748
120	14400	165	3778	3927

Table 3: Grid-refinement study for T2.

provide a basic sanity check that the method does not rely on an accidental grid-dependent parameter choice. In this small range of sizes, the cumulative inner work stays close to the outer count (about 4–5% more in Table 3), suggesting that the inexact Q-solve can be kept inexpensive when preconditioned by $M = P_{\mathrm{IC}}$. At the same time, the Rayleigh outer iterations grow markedly faster than PCG–ICscaled iterations, which is consistent with the fact that PCG is optimal in a Krylov subspace whereas the Newmark-induced iteration is a fixed-parameter stationary process. A detailed complexity study (wall time, matvec counts, and factorization cost) is beyond the scope of this note and is left for future work.

7 Conclusions

We developed a dynamical viewpoint in which explicit Newmark time stepping, applied to a suitably chosen *damped second-order dynamics*, induces stationary iterations for solving large sparse SPD linear systems $Au = b$. In this interpretation the "mass" operator M plays the role of a preconditioner, while the "damping" operator C introduces memory and spectral filtering, thus shaping the effective iteration polynomial.

From the algorithmic side, proportional damping ($C = \eta M$) yields inexpensive smoother-like schemes that can stagnate at moderate accuracies, whereas Rayleigh damping ($C = a_0 M + a_1 A$), combined with an effective choice of M (e.g., IC-scaled approximations), produces robust solver behavior. In our diffusion benchmarks, including high-contrast heterogeneous coefficients, the Rayleigh design with $M = P_{\mathrm{IC}}$ reaches tight tolerances using an inexact inner solve on $Q = M + \gamma\tau C$ whose total inner PCG iterations remain comparable to the number of outer steps.

Sensitivity tests indicate that moderate-to-high damping ratios (e.g. $\zeta \approx 0.7$–0.9 in the reported setting) offer a good trade-off between contraction and work.

Several directions deserve further investigation. While we used a constant time step τ, Newmark naturally allows variable τ_k; designing τ_k adaptively may provide a non-stationary acceleration akin to Chebyshev-type strategies and further reduce iteration counts on difficult spectra. Automated parameter selection for (a_0, a_1) and ζ based on inexpensive spectral estimates or residual-driven criteria would also strengthen robustness. Finally, extending the approach beyond SPD systems and to block-structured problems is a natural next step.

Acknowledgments

The present research is part of the activities of the "Dipartimento di Eccellenza 2023-2027". The author is a member of the "Gruppi Nazionali di Ricerca INdAM".

A Modal analysis

Standing assumptions. In this appendix we assume that A and M are symmetric positive definite. Moreover, we assume *classical (proportional) damping*, namely that the damping matrix C is diagonalizable in the same M-orthonormal modal basis as the pencil (A, M), with strictly positive modal damping coefficients (equivalently, (49) holds with $D \succ 0$). Under these assumptions we prove (6) (global asymptotic stability in the sense of Lyapunov) for (5). Let $(\omega_i^2, \boldsymbol{\phi}_i)$ be a generalized eigenpair of the pencil (A, M), defined by

$$A\,\boldsymbol{\phi}_i = \omega_i^2 M\,\boldsymbol{\phi}_i, \quad i = 1, 2, \ldots, n, \tag{43}$$

where $\omega_i > 0$ and $\boldsymbol{\phi}_i \in \mathbb{R}^n$ is the i-th normal mode. These modes form a basis of $\mathbb{R}^n$ and are assumed to be orthonormal with respect to the mass matrix, i.e.,

$$\boldsymbol{\phi}_i^T\, M\, \boldsymbol{\phi}_j = \delta_{ij}, \tag{44}$$

where δ_{ij} is the Kronecker symbol. In matrix form, (43) reads

$$A\,\Phi = M\,\Phi\,\Omega^2, \tag{45}$$

where $\Phi, \Omega^2 \in \mathbb{R}^{n\times n}$ are defined by $\Phi = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_n]$ and $\Omega^2 = \mathrm{diag}(\omega_i^2)$.
For any time t, let

$$\mathbf{w}(t) = \Phi\,\mathbf{y}(t) \tag{46}$$

be the expansion of the solution of (5) in terms of the normal modes. Substituting (46) into (5) yields

$$\begin{cases} M\Phi\,\ddot{\mathbf{y}} + C\Phi\,\dot{\mathbf{y}} + A\Phi\,\mathbf{y} = 0, \quad t > 0 \\ \Phi\,\mathbf{y}(0) = \Phi\,\mathbf{y}^{(0)} \\ \Phi\,\dot{\mathbf{y}}(0) = \Phi\,\mathbf{y}_1^{(0)}, \end{cases} \tag{47}$$

with $\mathbf{w}^{(0)} = \Phi\,\mathbf{y}^{(0)}$ and $\mathbf{w}_1^{(0)} = \Phi\,\mathbf{y}_1^{(0)}$. Using (45) in the first equation of (47), multiplying on the left by $(M\Phi)^{-1}$, and applying Φ^{-1} to the initial conditions, we obtain

$$\begin{cases} \ddot{\mathbf{y}} + \Phi^{-1}M^{-1}C\Phi\,\dot{\mathbf{y}} + \Omega^2\,\mathbf{y} = 0, \quad t > 0 \\ \mathbf{y}(0) = \mathbf{y}^{(0)} \\ \dot{\mathbf{y}}(0) = \mathbf{y}_1^{(0)}. \end{cases} \tag{48}$$

In the undamped case ($C = 0$), (48) decouples into n independent second-order equations since Ω^2 is diagonal. This is no longer true when $C \neq 0$. To still obtain a decoupled system we impose the condition

$$\Phi^{-1}M^{-1}C\Phi = D, \tag{49}$$

for some diagonal matrix $D \in \mathbb{R}^{n\times n}$ with positive entries. Equivalently,

$$C\,\Phi = M\,\Phi\,D, \tag{50}$$

i.e., the normal modes of (45) are also eigenvectors of the pencil (C, M). A sufficient (and, in this setting, essentially equivalent) condition is that A and C commute[1], namely

$$CM^{-1}A = AM^{-1}C. \tag{51}$$

Two classical choices used to enforce (49) (and hence (51)) are Rayleigh and Caughey damping. In the Rayleigh model,

$$C = a_0\,M + a_1\,A, \tag{52}$$

for some positive constants a_0, a_1. The Caughey model reads

$$C = M\sum_{i=-\infty}^{+\infty} a_i\,(M^{-1}\,A)^i. \tag{53}$$

Rayleigh damping is recovered as the special case $a_i = 0$ except for $i = 0$ and $i = 1$. Moreover, by the Cayley–Hamilton theorem applied to $M^{-1}A$, the number of linearly independent terms in (53) is at most n.

When (49) holds, system (48) decouples into the scalar equations

$$\begin{cases} \ddot{y}_i + 2\,\xi_i\,\omega_i\,\dot{y}_i + \omega_i^2\,y_i = 0, \quad t > 0 \\ y_i(0) = y_i^{(0)} \\ \dot{y}_i(0) = y_{1,i}^{(0)}, \end{cases} \tag{54}$$

[1]This commutativity condition (equivalently, simultaneous diagonalization by the undamped modal basis) goes back to Caughey [12] and to the necessary-and-sufficient characterization of Caughey and O'Kelly [13]; see also Adhikari and Phani [14] for a concise modern discussion.

for $i = 1, \dots, n$. Here the diagonal entries $d_{ii} > 0$ of D in (50) are written as

$$d_{ii} = 2\,\xi_i\,\omega_i, \tag{55}$$

where $\xi_i > 0$ is the damping ratio of the i-th mode. The solutions of (54) are linear combinations of

$$\begin{aligned} & e^{-\xi_i\omega_i t}\Big(\alpha_i \cos\big(\omega_i\sqrt{1-\xi_i^2}\,t\big) + \beta_i \sin\big(\omega_i\sqrt{1-\xi_i^2}\,t\big)\Big) && \xi_i < 1 \\ & e^{-\xi_i\omega_i t}\Big(\alpha_i e^{\omega_i\sqrt{\xi_i^2-1}\,t} + \beta_i e^{-\omega_i\sqrt{\xi_i^2-1}\,t}\Big) && \xi_i > 1 \\ & e^{-\xi_i\omega_i t}(\alpha_i + \beta_i\,t) && \xi_i = 1, \end{aligned} \tag{56}$$

corresponding to the underdamped, overdamped, and critically damped cases, respectively. In all cases $\lim_{t\to+\infty} y_i(t) = 0$, which implies (6) after returning to $\mathbf{w}$ via (46).

B State-space representation

Standing assumptions. In this appendix we assume that A and M are symmetric positive definite and that C is symmetric *positive definite*. (The argument can be adapted to $C \succeq 0$, but in that case one generally obtains Lyapunov stability rather than asymptotic decay unless all undamped modes are excluded.) Under these assumptions we give a second proof of (6) based on a state-space formulation. We rewrite (2) as a first-order time-invariant linear system, compute its solution via the matrix exponential, and infer (6) from the spectral properties of the associated matrix.

We proceed in two steps. First, introduce the change of variables

$$\mathbf{z} = M^{1/2}\,\mathbf{w}, \tag{57}$$

where $M^{1/2}$ denotes the (unique) symmetric positive definite square root of M. This yields

$$\begin{cases} \ddot{\mathbf{z}} + \underbrace{M^{-1/2}CM^{-1/2}}_{F}\,\dot{\mathbf{z}} + \underbrace{M^{-1/2}AM^{-1/2}}_{E^2}\,\mathbf{z} = 0, \quad t > 0 \\ \mathbf{z}(0) = M^{1/2}\,\mathbf{w}^{(0)} \\ \dot{\mathbf{z}}(0) = M^{1/2}\,\mathbf{w}_1^{(0)}, \end{cases} \tag{58}$$

with F and E symmetric positive definite. Next, introduce the auxiliary variables

$$\boldsymbol{\mathcal{U}} = E\,\mathbf{z}, \qquad \boldsymbol{\mathcal{V}} = \dot{\mathbf{z}}. \tag{59}$$

Using $(\boldsymbol{\mathcal{U}}, \boldsymbol{\mathcal{V}})$ instead of the more natural pair $(\mathbf{w}, \dot{\mathbf{w}})$ leads to a state-space representation for which the location of the eigenvalues is straightforward to characterize. With (59), system (58) becomes

$$\begin{cases} \dot{X} + \mathcal{A}\,X = 0, \quad t > 0 \\ X(0) = X_0, \end{cases} \tag{60}$$

where

$$X = \begin{pmatrix} \boldsymbol{\mathcal{U}} \\ \boldsymbol{\mathcal{V}} \end{pmatrix}, \quad \mathcal{A} = \begin{pmatrix} 0 & -E \\ E & F \end{pmatrix}, \quad X_0 = \begin{pmatrix} EM^{1/2}\,\mathbf{w}^{(0)} \\ M^{1/2}\mathbf{w}_1^{(0)} \end{pmatrix}. \tag{61}$$

The solution of (60) is given by the matrix exponential

$$X(t) = e^{-\mathcal{A}\,t}\,X_0. \tag{62}$$

The key point is that the eigenvalues of $\mathcal{A}$ lie in the open right half-plane.

Lemma B.1. *The eigenvalues of $\mathcal{A}$ have strictly positive real parts.*

Proof. Let (λ, Y) be an eigenpair of $\mathcal{A}$ (with possibly complex entries). Then

$$Y^*\,\mathcal{A}\,Y = \lambda\,\|Y\|^2,$$

where * denotes the Hermitian conjugate and $\|Y\|^2 = Y^*Y$. Writing $Y = (\mathbf{y}_1^T\,\mathbf{y}_2^T)^T$, a direct computation gives

$$Y^*\,\mathcal{A}\,Y = \underbrace{\mathbf{y}_2^*\,F\,\mathbf{y}_2}_{T_1} + \underbrace{\mathbf{y}_2^*\,E\,\mathbf{y}_1 - \mathbf{y}_1^*\,E\,\mathbf{y}_2}_{T_2}.$$

Taking the complex conjugate, we obtain

$$(Y^*\,\mathcal{A}\,Y)^* = \underbrace{\mathbf{y}_2^*\,F\,\mathbf{y}_2}_{T_1} + \underbrace{\mathbf{y}_1^*\,E\,\mathbf{y}_2 - \mathbf{y}_2^*\,E\,\mathbf{y}_1}_{-T_2}.$$

Hence T_1 is real and T_2 is purely imaginary. Therefore,

$$\Re\,(Y^*\,\mathcal{A}\,Y) = \mathbf{y}_2^*\,F\,\mathbf{y}_2 = \Re\,(\lambda)\,\|Y\|^2, \tag{63}$$

and since F is positive definite we deduce $\Re(\lambda) \geq 0$. Moreover, the eigenvalue relation $\mathcal{A}Y = \lambda Y$ is equivalent to

$$\begin{cases} \phantom{E\,\mathbf{y}_1} - \phantom{F\,\mathbf{y}_2} & E\,\mathbf{y}_2 = \lambda\,\mathbf{y}_1 \\ E\,\mathbf{y}_1 + F\,\mathbf{y}_2 = \lambda\,\mathbf{y}_2. \end{cases} \tag{64}$$

If $\Re(\lambda) = 0$, then (63) implies $\mathbf{y}_2 = 0$, and $(64)_2$ yields $E\mathbf{y}_1 = 0$. Since E is nonsingular, we get $\mathbf{y}_1 = 0$, contradicting $Y \neq 0$. Hence $\Re(\lambda) > 0$. (In particular, the same argument shows that $\mathcal{A}$ is nonsingular.) □

Lemma B.1 implies that the state-space system (60) is globally asymptotically stable, and therefore $\lim_{t\to+\infty} X(t) = 0$ for all X_0. Finally, (6) follows by returning to the primitive variable $\mathbf{w}$ through (60), (59), and (57).

References

[1] N. M. Newmark. A method of computation for structural dynamics. *Journal of the Engineering Mechanics Division (ASCE)*, 85(EM3):67–94, 1959.

[2] K.-J. Bathe. *Finite Element Procedures*. Prentice Hall, 1996.

[3] H. P. Gavin. *Numerical integration in structural dynamics* (lecture notes), Duke University, 2020.

[4] B. T. Polyak. Some methods of speeding up the convergence of iteration methods. *USSR Computational Mathematics and Mathematical Physics*, 4(5):1–17, 1964.

[5] R. S. Varga. A comparison of the successive overrelaxation method and semi-iterative methods using Chebyshev polynomials. *Journal of the Society for Industrial and Applied Mathematics*, 5(2):39–46, 1957.

[6] M. R. Hestenes and E. Stiefel. Methods of conjugate gradients for solving linear systems. *Journal of Research of the National Bureau of Standards*, 49:409–436, 1952.

[7] A. Greenbaum. *Iterative Methods for Solving Linear Systems*. SIAM, 1997.

[8] Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, 2nd edition, 2003.

[9] O. Axelsson. *Iterative Solution Methods*. Cambridge University Press, 1994.

[10] M. Benzi. Preconditioning techniques for large linear systems: A survey. *Journal of Computational Physics*, 182:418–477, 2002.

[11] E. I. Jury. *Theory and Application of the Z-Transform Method*. Wiley, 1964.

[12] T. K. Caughey. Classical normal modes in damped linear dynamic systems. *Journal of Applied Mechanics (ASME)*, 27:269–271, 1960.

[13] T. K. Caughey and M. E. J. O'Kelly. Classical normal modes in damped linear dynamic systems. *Journal of Applied Mechanics (ASME)*, 32(3):583–588, 1965.

[14] S. Adhikari and A. S. Phani. Rayleigh's classical damping revisited. In *International Conference on Civil Engineering in the New Millennium: Opportunities and Challenges*, Howrah, India, January 2007.

[15] A. S. Day. An introduction to dynamic relaxation. *The Engineer*, 219:218–221, 1965.

[16] J. R. H. Otter, A. C. Cassell, and R. E. Hobbs. Dynamic relaxation. *Proceedings of the Institution of Civil Engineers*, 35(4):633–656, 1966.

[17] P. Underwood. Dynamic relaxation. In T. Belytschko and T. J. R. Hughes, editors, *Computational Methods for Transient Analysis*, pages 245–265. North-Holland, 1983.

[18] C. T. Kelley and D. E. Keyes. Convergence analysis of pseudo-transient continuation. *SIAM Journal on Numerical Analysis*, 35(2):508–523, 1998.

[19] T. S. Coffey, C. T. Kelley, and D. E. Keyes. Pseudotransient continuation and differential-algebraic equations. *SIAM Journal on Scientific Computing*, 25(2):553–569, 2003.

[20] Y. Saad and H. A. van der Vorst. Iterative solution of linear systems in the 20th century. *Journal of Computational and Applied Mathematics*, 123(1–2):1–33, 2000.

MOX Technical Reports, last issues

Dipartimento di Matematica
Politecnico di Milano, Via Bonardi 9 - 20133 Milano (Italy)

Micheletti, S.
Newmark time marching as a preconditioned iteration for large SPD linear systems

07/2026 Corti, M.; Ahern, A.; Goriely, A.; Kuhl, E.; Antonietti, P.F.
A whole-brain model of amyloid beta accumulation and cerebral hypoperfusion in Alzheimer's disease

06/2026 Corti, M.; Gómez, S.
On the compact discontinuous Galerkin method for polytopal meshes

05/2026 Ranno, A.; Ballarin, F.; Lespagnol, F.; Zunino, P.; Perotto, S.
A fictitious domain formulation based on hierarchical model reduction applied to drug-eluting stents

04/2026 Ragni, A.; Cavinato, L.; Ieva, F.
Penalized Likelihood Optimization for Adaptive Neighborhood Clustering in Time-to-Event Data with Group-Level Heterogeneity

03/2026 Mapelli, A.; Carini, L.; Ieva, F.; Sommariva, S.
A neighbour selection approach for identifying differential networks in conditional functional graphical models

02/2026 Antonietti, P.F.; Beirao da Veiga, L.; Botti, M.; Harnist, A.; Vacca, G.; Verani, M.
Virtual Element methods for non-Newtonian shear-thickening fluid flow problems

01/2026 Iapaolo V.; Vergani, A.M.; Cavinato, L.; Ieva, F.
Multi-view learning and omics integration: a unified perspective with applications to healthcare

Iapaolo, V.; Vergani, A.; Cavinato, L.; Ieva, F.
Multi-view learning and omics integration: a unified perspective with applications to healthcare

82/2025 Varetti, E.; Torzoni, M.; Tezzele, M.; Manzoni, A.
Adaptive digital twins for predictive decision-making: Online Bayesian learning of transition dynamics