# Proof of the Forsythe Conjecture for $s = 2$

Jarek Liesen

July 30, 2026

## 0 Foreword and Method

The proof that follows was found on 26 July, 2026 using Codex with GPT 5.6-Sol at Ultra reasoning, and is currently undergoing expert review. Early experiments immediately revealed two major failure modes.

Firstly, agents easily got lost in computational experiments. The increments $\|v_{k+2} - v_k\|$ become small quickly, so numerical experiments cannot indicate convergence. Alternatively, the agents ran dozens of symbolic experiments, attempting to find particular families for which convergence could be established. We therefore decided to forbid experiments during the search for the proof and solely rely on symbolic manipulation.

Secondly, agents reliably drifted away from the original notation, often introducing unnecessary complexity. This complexity often obscured insights that would have been obvious with notation used by experts. Therefore, we first fixed the notation by running an audit of existing papers [1, 3] to generate a specification of the common mathematical language in this area. We subsequently used this audit to anchor agents during their search.

We instructed agents to create independent notes on intermediate lemmas during their proof search. Each document communicating a lemma was independently reviewed by two subagents for errors and notation quality. We found the following two strategies particularly effective: explicitly claim that there is an error in the proof, and ask if the generated document was written by the same authors as an expert paper [1]. The main agent was instructed to iterate on this feedback until the reviewers found no major issues.

Using this strategy, the first valid proof was found after 2 hours and 5 minutes. It is based on a repair of a previous proof attempt of Zhuk–Bondarenko [4], which is known to be incomplete. Subsequently, we searched for a substantial simplification, instructing the agent to distil the main arguments and required mathematical properties. Within 50 minutes, the agent discovered a three-step Arnoldi cancellation that establishes convergence of the polynomial factors directly, eliminating the need for the Zhuk–Bondarenko coefficient-convergence result and much of the machinery used in the first proof.

The simplified proof was then subjected to several adversarial reviews, including ChatGPT 5.6-Sol with Pro reasoning and Claude Fable 5 at max reasoning. Additionally, we asked Fable to rederive the proof from a short description of the proof strategy that did not include any results. It successfully reproduced the full proof within 8 minutes. Only then did we hand the proof off to an expert review, which is currently in progress.

The proof that follows was written by Codex. We are aware of a few minor issues, such as an inaccuracy in the equation following (10), but deliberately kept the proof untouched.

# 1 Introduction

The conjecture of Forsythe from 1968 [2] concerns the limiting directions in the conjugate gradient method restarted after a fixed number of steps. A formulation based on the Arnoldi projection was given in [1]; it applies to every real symmetric matrix. We use the notation of [3].

Let $A = A^T \in \mathbb{R}^{n\times n}$, let $v_0 \in \mathbb{R}^n$ satisfy $\|v_0\| = 1$, and suppose that $d(A, v_0) \geq 3$, where $d(A, v)$ denotes the grade of $v$ with respect to $A$. Put $\mathcal{K}_2(A, v) := \operatorname{span}\{v, Av\}$. For $k = 0, 1, 2, \ldots$, consider

$$\widetilde{v}_{k+1} = P_2(A; v_k)v_k, \tag{1}$$

$$v_{k+1} = \frac{\widetilde{v}_{k+1}}{\|\widetilde{v}_{k+1}\|}, \tag{2}$$

where $P_2(z; v_k)$ is the uniquely determined monic polynomial of degree two for which

$$\widetilde{v}_{k+1} \in A^2 v_k + \mathcal{K}_2(A, v_k) \quad \text{and} \quad \widetilde{v}_{k+1} \perp \mathcal{K}_2(A, v_k). \tag{3}$$

**Theorem 1.1.** *Let $A = A^T$, let $\|v_0\| = 1$, and suppose that $d(A, v_0) \geq 3$. Then the sequence $\{v_{2k}\}$ generated by (1)–(2) converges. Consequently, $\{v_{2k+1}\}$ also converges.*

We first derive convergence of the two-step polynomials directly from (3). The resulting limit equation reduces the proof to three or four spectral subspaces. With four subspaces, cubic interpolation removes the exterior components. An identity between consecutive two-step polynomials then proves convergence.

# 2 Proof of Theorem 1.1

By [1, Theorem 1], the iteration is well defined and $d(A, v_k) \geq 3$ for every $k$. Applying the iteration twice, we obtain

$$v_{k+2} = \frac{Q_k(A)v_k}{\|\widetilde{v}_{k+2}\|\,\|\widetilde{v}_{k+1}\|},$$

where $Q_k(z) := P_2(z; v_{k+1})P_2(z; v_k)$. For arbitrary real polynomials $f$ and $g$, we further write

$$\langle f, g\rangle_k := (g(A)v_k)^T f(A)v_k.$$

On polynomials of degree smaller than $d(A, v_k)$, this form is an inner product. Write $P_k(z) := P_2(z; v_k)$. The Arnoldi orthogonality conditions are

$$\langle P_k, 1\rangle_k = \langle P_k, z\rangle_k = 0. \tag{4}$$

Set $H_k := \|\widetilde{v}_{k+1}\|^2$. The change of vector then gives

$$\begin{aligned} \langle f, g\rangle_{k+1} &= \frac{\langle P_k f, P_k g\rangle_k}{H_k}, \\ \langle f, g\rangle_{k+2} &= \frac{\langle Q_k f, Q_k g\rangle_k}{H_k H_{k+1}}. \end{aligned} \tag{5}$$

Since $P_{k+1} - P_k$ has degree at most one, (4) and (5) give

$$\langle Q_k, 1\rangle_k = \langle P_k, P_{k+1}\rangle_k = H_k, \qquad \langle Q_k, Q_k\rangle_k = H_k H_{k+1}.$$

Set $R_k := Q_k - H_k$. The preceding identities give

$$\langle R_k, R_k \rangle_k = H_k(H_{k+1} - H_k). \tag{6}$$

In particular, $H_k$ is nondecreasing. The monic competitor $z^2$ gives

$$H_k \le \|A^2 v_k\|^2 \le \|A\|^4.$$

Hence

$$H_k \uparrow H > 0, \qquad \sum_{k=0}^{\infty}(H_{k+1} - H_k) = H - H_0.$$

The first identity above also yields

$$v_k^T v_{k+2} = \sqrt{\frac{H_k}{H_{k+1}}}. \tag{7}$$

If $H_{k+1} = H_k$, then (6) implies $R_k(A)v_k = 0$, and hence $v_{k+2} = v_k$. Uniqueness of the iteration makes it two-periodic from that index.

Every limit vector has grade at least three. Indeed, if $v_{k_j} \to v_*$ and $d(A, v_*) \le 2$, choose a monic quadratic $r$ such that $r(A)v_* = 0$. Since $P_{k_j}$ minimizes the norm among monic quadratics,

$$0 < H_0 \le H_{k_j} \le \|r(A)v_{k_j}\|^2 \longrightarrow 0,$$

a contradiction. Compactness now shows that the Gram matrices of $1, z$ in the forms $\langle \cdot, \cdot \rangle_k$ are uniformly positive definite. The two equations in (4) therefore give a uniform coefficient bound for $P_k$.We use $\| \cdot \|_{\mathrm{c}}$ for any fixed norm on the coefficients of polynomials of degree at most three.

**Lemma 2.1** (Factor variation)**.** *There is a constant $C > 0$, independent of $k$, such that*

$$\|P_{k+2} - P_k\|_{\mathrm{c}} \le C(H_{k+1} - H_k). \tag{8}$$

*Proof.* For every polynomial $\ell$ of degree at most one,

$$\langle P_k \ell, R_k \rangle_k = 0. \tag{9}$$

On the one hand,

$$\langle P_k \ell, Q_k \rangle_k = H_k \langle \ell, P_{k+1} \rangle_{k+1} = 0,$$

and, on the other hand,

$$\langle P_k \ell, H_k \rangle_k = H_k \langle P_k, \ell \rangle_k = 0$$

by (4). Subtracting the two equalities gives (9).

Let $D_k = P_{k+2} - P_k$, which has degree at most one. Orthogonality at step $k+2$ gives

$$\langle D_k, \ell \rangle_{k+2} = - \langle P_k, \ell \rangle_{k+2}.$$

Using (5) on the right-hand side, and expanding $Q_k^2 = (H_k + R_k)^2$, we see that the constant term vanishes by (4) and the term linear in $R_k$ vanishes by (9). We have therefore proved, for every such $\ell$,

$$\langle D_k, \ell \rangle_{k+2} = -\frac{\langle P_k \ell, R_k^2 \rangle_k}{H_k H_{k+1}}. \tag{10}$$

Taking $\ell = D_k$ in (10), using (6), and evaluating on the fixed finite spectrum of $A$, we can now compare the two sides:

$$c\|D_k\|_{\mathrm{c}}^2 \le \langle D_k, D_k \rangle_{k+2} = \frac{|\langle P_k D_k, R_k^2 \rangle_k|}{H_k H_{k+1}} \le C\|D_k\|_{\mathrm{c}}(H_{k+1} - H_k)$$

with constants $c, C > 0$. This proves (8). $\square$

*Proof of Theorem 1.1.* If $H_{k+1} = H_k$ for some $k$, then both parity subsequences are eventually constant. We may therefore suppose that

$$H_k < H \qquad \text{for every } k.$$

The summability of the differences $H_{k+1} - H_k$ and (8) show that the two parity subsequences converge coefficientwise to monic quadratics $p$ and $q$:

$$P_{2k} \longrightarrow p, \qquad P_{2k+1} \longrightarrow q. \tag{11}$$

Since $Q_k = P_{k+1}P_k$, it follows that $Q_k \to pq$. Write $Q := pq$. Moreover,

$$\|Q_k - Q\|_{\mathrm{c}} \le C(H - H_k). \tag{12}$$

Indeed, the two parity subsequences of $P_k$ have finite total variation. Finally,

$$Q_{k+1} - Q_k = P_{k+1}(P_{k+2} - P_k),$$

so the uniform coefficient bound, (8), and telescoping from $k$ to infinity give (12).

Let $\Sigma_A^*$ denote the set of limit vectors of $\{v_{2k}\}$. It is compact. Moreover, (7) gives

$$\|v_{k+2} - v_k\|^2 = 2\left(1 - \sqrt{\frac{H_k}{H_{k+1}}}\right) \longrightarrow 0.$$

Thus each parity limit set is connected: if it were the disjoint union of two nonempty compact sets, its tail would lie in two separated neighborhoods and visit both infinitely often, which is impossible when successive parity increments tend to zero.

If $v_{2k_j} \to v_*$, then (7) gives $v_{2k_j+2} \to v_*$. Passing to the limit in the two-step recurrence gives

$$(Q(A) - HI)v_* = 0. \tag{13}$$

The same holds for odd limit vectors. The known lower grade bound proven above and the degree of $Q - H$ now give

$$3 \le d(A, v_*) \le 4$$

for every even or odd limit vector $v_*$.

Choose one unit vector $u_j$ in the direction of the projection of $v_0$ onto each initially active eigenspace, and let $\mu_j$ be its eigenvalue. Write

$$v_k = \sum_j \nu_{k,j} u_j.$$

A vanished component remains zero. Each of the finitely many initially active components can therefore be deleted at most once. Choose an even index after the last deletion and restrict the iteration to the tail beginning there. This does not change either parity limit set. We may therefore suppose that

$$\nu_{k,j} \neq 0$$

for every remaining active eigenvalue and every $k$.

Among the remaining active eigenvalues, let

$$I_+ = \{\mu_j : Q(\mu_j) = H\}.$$

By (13), every even or odd limit vector is supported on $I_+$. Since every limit vector has grade at least three and $Q-H$ is monic of degree four,

$$3 \le |I_+| \le 4.$$

On any prescribed three-node support $\{\lambda_1, \lambda_2, \lambda_3\} \subseteq I_+$, the squared coefficients $w_i$ of an even limit vector satisfy

$$\sum_i w_i = 1, \qquad \sum_i w_i p(\lambda_i) = 0, \qquad \sum_i w_i \lambda_i p(\lambda_i) = 0.$$

These equations have a unique solution. Indeed, $p(\lambda_i)q(\lambda_i) = H$, and after factoring $p(\lambda_i)$ from the $i$-th column, the determinant becomes the evaluation determinant of $q/H, 1, z$ at three distinct nodes, which is nonzero. Moreover, at every $\lambda_i \in I_+$ the two-step multiplier tends to $Q(\lambda_i)/H = 1$, so its sign is eventually positive on each parity. Hence $|I_+| = 3$ gives convergence of the even sequence, and then of the odd sequence by one application of (1)–(2).

It remains to consider

$$I_+ = S = \{\lambda_1 < \lambda_2 < \lambda_3 < \lambda_4\}.$$

For the corresponding eigendirections, relabel $u_j$ and $\nu_{k,j}$ as $u_i$ and $\nu_{k,i}$, respectively. We write

$$\Pi(z) = \prod_{i=1}^{4} (z - \lambda_i).$$

The four roots of $Q-H$ give

$$Q(z) = \Pi(z) + H = p(z)q(z). \tag{14}$$

If $\Sigma_A^*$ is a singleton, the theorem follows as in the three-node case. Suppose that it is not a singleton. It then contains a vector of grade four: otherwise the preceding calculation and the eventually fixed signs would make the connected set $\Sigma_A^*$ finite.

The limit equation shows only that the components outside $S$ tend to zero. To treat their influence on the four principal components as a summable perturbation, we need the stronger estimate in the next lemma.

**Lemma 2.2** (Exterior decay)**.** *Assume that $S = I_+$ consists of four points and that $\Sigma_A^*$ contains a vector of grade four. Then there are constants $C > 0$ and $0 < \vartheta < 1$ such that*

$$\varepsilon_k := \sum_{\mu_j \notin S} \nu_{k,j}^2 \le C \vartheta^k. \tag{15}$$

*Proof.* Let $\mu_j \notin S$ be a remaining active eigenvalue. Its exact two-step recurrence is

$$\nu_{k+2,j} = \frac{Q_k(\mu_j)}{\sqrt{H_k H_{k+1}}} \nu_{k,j}. \tag{16}$$

Every such component tends to zero on both parities by (13). It follows from (16) that $|Q(\mu_j)| \le H$. If the inequality is strict, the absolute value of the multiplier in (16) is bounded by a number smaller than one on each sufficiently late parity, and the component decreases geometrically. Equality with $Q(\mu_j) = H$ is excluded by the definition of $S$. It remains to exclude $Q(\mu_j) = -H$.

Assume this equality and define

$$\beta_{i,j} = \frac{\Pi(\mu_j)}{(\mu_j - \lambda_i)\Pi'(\lambda_i)}, \qquad i = 1, \ldots, 4.$$

These are the four cubic Lagrange polynomials evaluated at $\mu_j$. Consequently,

$$\sum_i \beta_{i,j} = 1, \qquad F(\mu_j) = \sum_i \beta_{i,j} F(\lambda_i) \quad \text{for every } \deg F \le 3.$$

Define

$$Z_{j,k} = \nu_{k,j}^2 \prod_{i=1}^{4} (\nu_{k,i}^2)^{\beta_{i,j}}.$$

All factors are positive because the last deletion has already been discarded. Moreover, $Q_k(\lambda_i) > 0$ for all sufficiently large $k$, since $Q_k(\lambda_i) \to H$. Applying (16) at $\mu_j$ and at the four nodes, with the real powers interpreted through logarithms, gives

$$\frac{Z_{j,k+2}}{Z_{j,k}} = \left( \frac{|Q_k(\mu_j)| \prod_i Q_k(\lambda_i)^{\beta_{i,j}}}{H_k H_{k+1}} \right)^2.$$

Write $h_k := H - H_k$. By (12), $Q_k - Q = O(h_k)$ coefficientwise, and $\deg(Q_k - Q) \le 3$. Cubic interpolation cancels the complete first variation of the numerator, and therefore

$$\log \frac{|Q_k(\mu_j)| \prod_i Q_k(\lambda_i)^{\beta_{i,j}}}{H_k H_{k+1}} = \frac{h_k + h_{k+1}}{H} + O(h_k^2) > 0$$

for all sufficiently large $k$; here $0 < h_{k+1} \le h_k$, so the linear term dominates the remainder.

Along an even subsequence converging to a grade-four point, the four numbers $\nu_{k,i}^2$, $i = 1, \ldots, 4$, converge to positive values, whereas $\nu_{k,j}^2 \to 0$. Thus $Z_{j,k} \to 0$ along this subsequence, contrary to its eventual increase on the even indices. Hence the equality $Q(\mu_j) = -H$ is impossible. There are only finitely many exterior eigenvalues and two parities. Taking the largest of their finitely many contraction rates gives (15). □

The lemma leaves only four principal eigendirections. The remaining possible obstruction is drift of their squared coefficients even though the polynomials already converge. Write $w_{k,i} := \nu_{k,i}^2$, and normalize these weights by

$$x_{k,i} = \frac{w_{k,i}}{1 - \varepsilon_k}, \qquad i = 1, \ldots, 4. \tag{17}$$

Let $\Delta_3$ denote the nonnegative vectors $x = (x_1, \ldots, x_4)^T$ with $\sum_i x_i = 1$ and at least three positive entries. For $x \in \Delta_3$, let $P_x$ be the monic quadratic determined by

$$\sum_i x_i P_x(\lambda_i) = \sum_i x_i \lambda_i P_x(\lambda_i) = 0,$$

and put

$$\mathcal{H}(x) = \sum_i x_i P_x(\lambda_i)^2, \qquad \mathcal{T}(x)_i = \frac{x_i P_x(\lambda_i)^2}{\mathcal{H}(x)}.$$

Thus $P_x$ is the one-step polynomial for the weights $x$, $\mathcal{H}(x)$ is its squared norm, and $\mathcal{T}(x)$ gives the next squared weights. Their defining formulas are locally smoothly extendible at every point of $\Delta_3$, and $\mathcal{T}(x) \in \Delta_3$.

Put $E_i := \Pi'(\lambda_i)$. The two orthogonality equations say that the vector with entries $x_i P_x(\lambda_i)$ is orthogonal to the vectors with entries 1 and $\lambda_i$. The four-node Lagrange identities therefore give constants $a, b$ such that

$$x_i P_x(\lambda_i) = \frac{a\lambda_i + b}{E_i}.$$

Since $P_x$ is monic of degree two, the same identities give $a = \mathcal{H}(x) > 0$. Thus there is a number $\rho(x) = -b/a$ such that

$$x_i P_x(\lambda_i) = \mathcal{H}(x)\frac{\lambda_i - \rho(x)}{E_i}. \tag{18}$$

For every fixed $i$,

$$\rho(x) = \lambda_i - \frac{E_i x_i P_x(\lambda_i)}{\mathcal{H}(x)},$$

so $\rho$ is smooth also at a three-node boundary state.

Every even limit weight has polynomial $p$ and squared norm $H$. Equation (18) and the grade-four limit point therefore select a closed gap $J$ between two consecutive nodes such that the even and odd limit weights lie respectively on the curves $\Gamma_p$ and $\Gamma_q$ parametrized by

$$x_i(\rho) = H\frac{\lambda_i - \rho}{E_i p(\lambda_i)}, \quad y_i(\rho) = H\frac{\lambda_i - \rho}{E_i q(\lambda_i)}, \qquad \rho \in J. \tag{19}$$

Conversely, the Lagrange identities and (14) show that all the weights in (19) belong to these two curves. At an interior grade-four weight, the signs of $(\lambda_i - \rho)/(E_i p(\lambda_i))$ fix one of the three node gaps. Since $p(\lambda_i)q(\lambda_i) = H > 0$, the second formula has the same sign pattern and selects the same gap. That pattern places the two zeros of each polynomial in the other two node gaps, and hence

$$\min_{\rho \in J} |p(\rho)q(\rho)| > 0. \tag{20}$$

If $x \in \Gamma_p$ and $y = \mathcal{T}(x)$, then

$$y_i q(\lambda_i) = \frac{x_i p(\lambda_i)^2 q(\lambda_i)}{H} = x_i p(\lambda_i).$$

Thus $P_y = q$, and

$$\mathcal{H}(y) = \frac{1}{H}\sum_i x_i p(\lambda_i)^2 q(\lambda_i)^2 = H.$$

Hence $y \in \Gamma_q$. The same calculation with $p, q$ interchanged shows that $\mathcal{T}(\Gamma_q) \subseteq \Gamma_p$. In particular,

$$P_x P_{\mathcal{T}(x)} = \Pi + H \qquad (x \in \Gamma_p \cup \Gamma_q). \tag{21}$$

Put $x_k := (x_{k,1}, \ldots, x_{k,4})^T$, the vector of normalized principal squared coefficients. Its actual update is close to this exact four-node map. Smooth dependence of the two orthogonality equations and (15) give

$$P_k = P_{x_k} + O(\varepsilon_k) \quad \text{coefficientwise}, \qquad H_k = \mathcal{H}(x_k) + O(\varepsilon_k).$$

Moreover,

$$x_{k+1,i} = \frac{x_{k,i} P_k(\lambda_i)^2}{\sum_{r=1}^4 x_{k,r} P_k(\lambda_r)^2}.$$

Consequently,

$$x_{k+1} = \mathcal{T}(x_k) + O(\vartheta^k). \tag{22}$$

**Lemma 2.3** (The four-node identity)**.** *Let $x \in \Delta_3$ and $y = \mathcal{T}(x)$. For $t \in \mathbb{R}$, put*

$$D_t(z) := \frac{\Pi(z) - \Pi(t)}{z - t}.$$

*For $u \in \Delta_3$, let $\alpha(u)$ be the coefficient of $z^3$ in $P_u P_{\mathcal{T}(u)} - \Pi$. Then*

$$\begin{aligned} P_x(z)P_y(z) &= \Pi(z) + \mathcal{H}(y) + \alpha(x)D_{\rho(x)}(z), \\ \rho(y) - \rho(x) &= \frac{\alpha(x)\Pi(\rho(x))}{\mathcal{H}(y)}. \end{aligned} \tag{23}$$

*For a monic quadratic $f$, let $\ell_f(t)$ be the coefficient of $z$ in the remainder of $D_t$ after division by $f$. Then*

$$\alpha(y)\ell_{P_y}(\rho(y)) = \alpha(x)\ell_{P_y}(\rho(x)), \tag{24}$$

*and, for every $t \in \mathbb{R}$, the limiting polynomials satisfy*

$$\ell_q(t) = p(t), \qquad \ell_p(t) = q(t). \tag{25}$$

*Proof.* Suppose first that both $x$ and $y$ have four positive entries. Applying (18) first to $x$ and then to $y$, we obtain at every node

$$P_x(\lambda_i)P_y(\lambda_i) = \mathcal{H}(y)\frac{\lambda_i - \rho(y)}{\lambda_i - \rho(x)}.$$

Let

$$B = P_x P_y - \Pi - \mathcal{H}(y).$$

On the one hand, $\deg B \leq 3$, and at the four nodes

$$B(\lambda_i) = -\mathcal{H}(y)\frac{\rho(y) - \rho(x)}{\lambda_i - \rho(x)}, \qquad D_{\rho(x)}(\lambda_i) = -\frac{\Pi(\rho(x))}{\lambda_i - \rho(x)}.$$

On the other hand, the coefficient of $z^3$ in $B$ is $\alpha(x)$. Comparing the four values and using uniqueness of cubic interpolation gives both identities in (23). States for which both $x$ and $y$ have four positive entries are dense in $\Delta_3$, and all terms extend continuously, so the remaining cases follow.

Two consecutive products in the first identity share the quadratic $P_y$. If $z = \mathcal{T}(y)$, reduction modulo $P_y$ gives

$$\begin{aligned} 0 &\equiv \Pi + \mathcal{H}(y) + \alpha(x)D_{\rho(x)} \qquad (\text{mod } P_y), \\ 0 &\equiv \Pi + \mathcal{H}(z) + \alpha(y)D_{\rho(y)} \qquad (\text{mod } P_y). \end{aligned}$$

The constants have zero coefficient of $z$. Comparing that coefficient proves (24). Finally, (14) gives

$$\frac{\Pi(z) - \Pi(t)}{z - t} = q(z)\frac{p(z) - p(t)}{z - t} + p(t)\frac{q(z) - q(t)}{z - t}.$$

Reduction modulo $q$ shows that the coefficient of $z$ in the remainder is $p(t)$. Interchanging $p$ and $q$ proves (25). □

*The perturbed sequence.* The quantity $\alpha(x)$ measures the failure of the exact two-step factorization $P_x P_{\mathcal{T}(x)} = \Pi + H$; by (21) it vanishes on $\Gamma_p \cup \Gamma_q$. The shared factor now turns the possible drift in $\rho$ into a scalar recurrence. We apply it directly to the weights $x_k$ in (17). Compactness and the description of their limit points imply

$$\operatorname{dist}(x_{2k}, \Gamma_p) \longrightarrow 0, \qquad \operatorname{dist}(x_{2k+1}, \Gamma_q) \longrightarrow 0.$$

Consequently,

$$P_{x_{2k}} \longrightarrow p, \qquad P_{x_{2k+1}} \longrightarrow q, \qquad \mathcal{H}(x_k) \longrightarrow H,$$

and $\rho(x_k)$ approaches $J$. Write $\rho_k := \rho(x_k)$. Put $\alpha_k := \alpha(x_k)$, and let $y_k := \mathcal{T}(x_k)$. The exact identities (23) and (24) give

$$\rho(y_k) - \rho_k = \frac{\Pi(\rho_k)}{\mathcal{H}(y_k)}\alpha_k,$$
$$\alpha(y_k)\ell_{P_{y_k}}(\rho(y_k)) = \alpha_k \ell_{P_{y_k}}(\rho_k).$$

All the maps in these formulas are smooth in a neighborhood of $\Gamma_p \cup \Gamma_q$. Comparing the exact point $y_k$ with the perturbed point $x_{k+1}$ in (22), we obtain, for all sufficiently large $k$,

$$\begin{aligned} \alpha_{k+1} &= m_k \alpha_k + O(\vartheta^k), \\ \rho_{k+1} - \rho_k &= c_k \alpha_k + O(\vartheta^k), \end{aligned} \tag{26}$$

where

$$m_k = \frac{\ell_{P_{y_k}}(\rho_k)}{\ell_{P_{y_k}}(\rho(y_k))}, \qquad c_k = \frac{\Pi(\rho_k)}{\mathcal{H}(y_k)}.$$

Since $\alpha = 0$ on $\Gamma_p \cup \Gamma_q$, we have $\alpha_k \to 0$. The first identity in (23) then gives $\rho(y_k) - \rho_k \to 0$. Equations (20) and (25) now show that the two factors defining $m_k$ have the same nonzero sign and that $m_k \to 1$.

Choose $K$ sufficiently large, let $M_K = 1$, and define

$$M_k = \prod_{j=K}^{k-1} m_j, \qquad k > K.$$

Choose $\epsilon > 0$ so that $\vartheta e^{2\epsilon} < 1$. After increasing $K$, for $j \geq k \geq K$,

$$e^{-\epsilon(j-k)} \leq \frac{M_j}{M_k} \leq e^{\epsilon(j-k)}.$$

Dividing the first relation in (26) by $M_{k+1}$ shows that $\alpha_k/M_k$ has a limit $L$, since its error increments are summable. If $L = 0$, summing those increments from infinity and using the displayed bounds gives $\alpha_k = O(\vartheta_1^k)$ for some $\vartheta_1 < 1$. The second relation in (26) then has summable increments.

If $L \neq 0$, then $\alpha_k$ is eventually one-signed and $\vartheta^k = o(|\alpha_k|)$. On every compact subinterval of the interior of $J$, $c_k$ has one sign and is bounded away from zero, so all late steps starting there have the same direction. Since $\rho_{k+1} - \rho_k \to 0$, two distinct limit points would enclose a compact subinterval of $J^\circ$ and force infinitely many crossings of it in both directions. Thus $\rho_k$ converges in either case.

Finally, (18), the convergence of $\rho_k$, the two polynomial limits above, and $\mathcal{H}(x_k) \to H$ show that each component of $x_k$ converges on each parity. Since $w_{k,i} = (1 - \varepsilon_k)x_{k,i}$, the four principal squared coefficients have the same limits. By (15), all exterior coefficients tend to zero. At every node $\lambda_i \in S$, the two-step multiplier in (16) tends to

$$\frac{Q(\lambda_i)}{H} = 1.$$

These multipliers are therefore positive for all sufficiently large $k$, so the signs of the four principal coefficients are fixed on the even subsequence. Thus $\{v_{2k}\}$ converges. Finally,

$$v_{2k+1} = \frac{P_{2k}(A)v_{2k}}{\sqrt{H_{2k}}},$$

and (11) gives convergence of the odd sequence. □

*Remark* 2.4. The coefficient limits (11) alone do not prove convergence of the vectors. In the four-node case, (19) describes a one-parameter family whose members have the same limiting polynomials $p, q$ and the same two-step polynomial $Q$. The shared-factor identity (24) supplies the additional information that controls motion within this family.

# References


[1] V. Faber, J. Liesen, and P. Tichý, *On the Forsythe conjecture*, BIT Numerical Mathematics, 63 (2023), Paper No. 49.

[2] G. E. Forsythe, *On the asymptotic directions of the s-dimensional optimum gradient method*, Numerische Mathematik, 11 (1968), pp. 57–76.

[3] J. Liesen, *A note on the restarted Arnoldi method for symmetric matrices*, working note (unpublished), July 3, 2026.

[4] P. F. Zhuk and L. N. Bondarenko, *A conjecture of G. E. Forsythe*, Mat. Sb. (N.S.), 121(163) (1983), no. 4, pp. 435–453; English translation, *On a conjecture of G. Forsythe*, Math. USSR-Sb., 49 (1984), no. 2, pp. 427–445.