```
Select a b ∈ ℂ^n, b ≠ 0
%% Initiation:
v = b/‖b‖_2, w = Av
V = [v], W = [w], H = [v*w]
repeat
     %% Extraction:
1)   Compute pre Ritz pairs (ϑ, y):   Hy = ϑy
2)   Selected an appropriate pre Ritz pair, say, (ϑ, y).
3)   u = Vy, r = Wy − ϑu
     %% Stopping criterion:
4)   if ‖r‖_2 < tol, x = u, λ = ϑ, break, end
     %% Expansion:
5)   Select an expansion vector t ∈ ℂ^n, t ≠ 0.
6)   [v, h⃗] = Orth(V, t)
7)   V ← [V, v], H ← [H; v*W]
8)   w = Av, W ← [W, w], H = [H, V*w]
end repeat
```

ALGORITHM 13.1. Subspace method for computing one eigenpair $(\lambda, \mathbf{x})$ with residual accuracy tol for a general square matrix $\mathbf{A}$.

March 26, 2018

Lecture 13 – Subspace Methods for Eigenvalue Problem

In this lecture, $\mathbf{A}$ is an $n \times n$ matrix.

A pseudo code for any subspace method for eigenvalue problems basically is as in ALG. 13.1. A (hopefully well-conditioned) basis $\mathbf{V}$ for a search space $\mathcal{V} = \text{span}(\mathbf{V})$ is **expanded** in each step (as in the substeps 5 – 8 of ALG. 13.1): an **expansion vector t** is selected (in substep 5) and is orthogonalised against $\mathbf{V}$ (in 6) before being included in the basis $\mathbf{V}$. An appropriate approximate eigenpair is **extracted** from the search subspace in each step (the substeps 1 – 3): H is the projection of $\mathbf{A}$ onto the search subspace $\mathcal{V}$. An appropriate solution (ϑ, y) of the projected problem $Hy = \vartheta y$ (in 1) is lifted to an approximate solution in n-space (in substep 3). Convergence is monitored (in 4): the stopping criterion.

Methods differ in the way the expansion vector is selected (see Lecture 13.E): better expansion vectors (ones that have a smaller angle with the required eigenvector) lead to faster convergence (less steps), but are usually harder to compute. The methods are named after the expansion strategy even though there are also different extraction strategies (see Lecture 13.A). Selection of the appropriate approximate eigenpair (see Lecture 13.B) depends on what eigenpairs are wanted.

Depending on the choices, more efficient implementation of substeps may be possible, for instance, with $\mathbf{t} = \mathbf{w} = \mathbf{Av}$, we retain Arnoldi's method and H can be expanded using $\vec{h}$ (cf., Exercise 13.1).

In practice, usually more then one eigenpair is required. Then, after detection of an eigenpair (cf., substep 4) the detected eigenvector has to be deflated from the matrix (see Lecture 13.D) before searching for the next eigenpair. If many steps are required, the search subspace $\mathcal{V}$ will be high dimensional, requiring a lot of storage space (to store $\mathbf{V}$, $\mathbf{W}$,...) and slows down the computational speed of the steps (substep 6 will become costly and possibly

also substep 1). Then restarts are required (see Lecture 13.C).

In substep 6, we suggested to orthogonalise the expansion vector against the previously computed basis vectors (the columns of $\mathbf{V}$). This is the most popular approach for obtaining a well-conditioned basis. However, there are methods that use a modified orthogonalisation procedure (as bi-Lanczos).

Exercise 13.1.

(a) Show that we obtain a method that is mathematically equivalent to Arnoldi's method if, in ALG. 13.1.5, we select $\mathbf{t} = \mathbf{w}(= \mathbf{A}\mathbf{v})$ (**Arnoldi's expansion**). Show that then the H that is formed in substep 7 equals $\underline{H}_k$ if $\mathbf{V}$ formed in substep 7 has $k+1$ columns. Simplify the computation of H in this case.

(b) Prove that in Arnoldi's method, in one step, the residuals of all Ritz pairs are co-linear (i.e., each residual is a scalar multiple of any other residual in that step).

(c) Show that we obtain a method that is mathematically equivalent to Arnoldi's method if, in ALG. 13.1.5, we select $\mathbf{t} = \mathbf{r}$ (**residual expansion**). Note that residual expansion does not require additional matrix-vector multiplications.

A Extraction

If a subspace $\mathcal{V}$ is available (in the form of a matrix $\mathbf{V}$ of basis vectors such that $\mathrm{span}(\mathbf{V}) = \mathcal{V}$), then the question is how to extract appropriate approximate eigenvectors from the space? For Krylov subspaces, two different approaches were suggested in Lecture 6, one with Ritz values, and another with harmonic Ritz values (see Exercise 6.4). Although harmonic Ritz values have been introduced in Lecture 6 mainly for their relation to GMRES (see Exercise 6.8), they certainly play a very useful role in eigenvalue computation as we will learn below.

We first recall the notion of Ritz values and harmonic Ritz values and generalise it to general spaces and general $n \times n$ matrices $\mathbf{A}$. A scalar $\vartheta \in \mathbb{C}$ and a non-trivial n-vector $\mathbf{u}$ forms a **Ritz pair** $(\vartheta, \mathbf{u})$ with **Ritz value** ϑ and **Ritz vector** $\mathbf{u}$ of $\mathbf{A}$ with respect to $\mathcal{V}$ if

$$\vartheta \in \mathbb{C}, \quad \mathbf{u} \in \mathcal{V}, \mathbf{u} \neq \mathbf{0} \quad \text{and} \quad \mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathbf{V},$$

$(\vartheta, \mathbf{u})$ is an **harmonic Ritz pair** with **harmonic Ritz value** ϑ and **harmonic Ritz vector** $\mathbf{u}$ of $\mathbf{A}$ with respect to $\mathcal{V}$ and target τ if

$$\vartheta \in \mathbb{C}, \quad \mathbf{u} \in \mathcal{V}, \mathbf{u} \neq \mathbf{0} \quad \text{and} \quad \mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp (\mathbf{A} - \tau\mathbf{I})\mathbf{V} :$$

for harmonic Ritz values, the test subspace is $\mathbf{A} - \tau\mathbf{I}$ times $\mathcal{V}$.[1] Here, τ is the **target value**, that is, we are interested in finding eigenvalues of $\mathbf{A}$ close to the scalar τ.

If $\mathcal{V}$ contains an good approximation of an eigenvector $\mathbf{x}$ with eigenvalue λ (that is, the angle between $\mathcal{V}$ and $\mathbf{x}$ is small), then there is a normalised vector $\widetilde{\mathbf{u}} = \mathbf{V}\widetilde{y}$ in $\mathcal{V}$, as the normalised orthogonal projection of $\mathbf{x}$ onto $\mathcal{V}$, for which the residual $\widetilde{\mathbf{r}} \equiv \mathbf{A}\widetilde{\mathbf{u}} - \widetilde{\vartheta}\widetilde{\mathbf{u}}$ is small. Here, $\widetilde{\vartheta}$ is the Rayleigh quotient, $\widetilde{\vartheta} \equiv \widetilde{\mathbf{u}}^*\mathbf{A}\widetilde{\mathbf{u}}$. With perturbation $\Delta \equiv -\widetilde{\mathbf{r}}\,\widetilde{\mathbf{u}}^*$, $(\widetilde{\vartheta}, \widetilde{\mathbf{u}})$ is an eigenpair of the perturbed problem $(\mathbf{A} + \Delta)\widetilde{\mathbf{u}} = \widetilde{\vartheta}\widetilde{\mathbf{u}}$. Since $\|\Delta\|_2 = \|\widetilde{\mathbf{r}}\|_2$, the perturbation Δ is small and $(\widetilde{\vartheta}, \widetilde{\mathbf{u}})$ is close to $(\lambda, \mathbf{x})$. Hence, $(\mathbf{V}^*\mathbf{A}\mathbf{V} + \mathbf{V}^*\Delta\mathbf{V})\widetilde{y} = \widetilde{\vartheta}\widetilde{y}$ with small perturbation $\mathbf{V}^*\Delta\mathbf{V}$: $\|\mathbf{V}^*\Delta\mathbf{V}\|_2 \leq \|\Delta\|_2 \leq \|\widetilde{\mathbf{r}}\|_2$. From the perturbation theorems of Lecture 12, we know that there is an eigenpair (ϑ, y) of $\mathbf{V}^*\mathbf{A}\mathbf{V}$ that is close to $(\widetilde{\vartheta}, \widetilde{y})$, that is, there is a Ritz pair $(\vartheta, \mathbf{u} \equiv \mathbf{V}y)$ that is close to $(\widetilde{\vartheta}, \widetilde{\mathbf{u}})$. Since this last pair is close to $(\lambda, \mathbf{x})$, we see that $(\vartheta, \mathbf{u})$ is close to $(\lambda, \mathbf{x})$. Moreover, the residual for $(\vartheta, \mathbf{u})$ is small. Summarising:

Proposition 13.1 *If $(\lambda, \mathbf{x})$ is a simple eigenpair of $\mathbf{A}$ and $\phi \equiv \angle(\mathcal{V}, \mathbf{x})$ is small, then there is a Ritz pair $(\vartheta, \mathbf{u})$ of $\mathbf{A}$ with respect to $\mathcal{V}$ that is (directionally) close to $(\lambda, \mathbf{x})$ and has small residual ('small' is of order ϕ).*

[1] We are enforcing a so-called Petrov–Galerkin condition. In Ritz–Galerkin test space and search space are the same. For this reason, harmonic Ritz values are also called Petrov values if $\mathbf{A}$ is non-normal.

In other words, if $\mathcal{V}$ contains a good approximation of an eigenvector, then a good approximation of that eigenvector can be detected by computing Ritz vectors and checking the sizes of their residuals. A similar analysis shows that also harmonic Ritz vectors can be used to detect good approximations of eigenvectors in situations where $\mathcal{V}$ contains good approximations of eigenvectors. Theorems as the ones in Lecture 12 can be used to quantify the above statements.

Exercise 13.2. Quantify (in terms of ϕ and $\tan(\phi)$) the above claims.

However, often appropriate approximate eigenpairs have to be identified when no small residuals exist. For instance, when the search subspace $\mathcal{V}$ is too large (high dimensional) and memory limitations or computational efficiency requires a 'restart' with a smaller (lower dimensional) linear subspace of $\mathcal{V}$. This will certainly occur if not only one, but a set of eigenpairs are to be computed. At restart, non of the residual may be small, or we want to include more approximate eigenpairs in the restart space than only the ones with small residual. Moreover, computing the size of the residuals of all approximate eigenpairs gives additional costs.

For Hermitian matrices, we saw several theorems (as Cauchy's interlace theorem II) that indicate that Ritz vectors may form appropriate approximate eigenvectors specifically in case the wanted eigenvalues are extremal (largest or smallest). Below, we will discuss an illustrative example. In case interior eigenvalues, more specifically, eigenvalues close to the target value τ in the interior of the spectrum, are to be computed, then harmonic Ritz vectors might be more appropriate as we will argue below. This claim follows from the observation that Harmonic Ritz vectors can be viewed as Ritz vectors for an inverted matrix:

Proposition 13.2 *Let* $\mathbf{u}$ *be an harmonic Ritz vector with respect to* $\mathcal{V}$ *and target* τ*.*
With $\mathbf{y} \equiv \mathbf{A}\mathbf{u} - \tau\mathbf{u}$ *and* $\mathbf{W} \equiv (\mathbf{A} - \tau\mathbf{I})\mathbf{V}$ *we have that* $(\frac{1}{\vartheta - \tau}, \mathbf{y})$ *is a Ritz pair of* $(\mathbf{A} - \tau\mathbf{I})^{-1}$ *with respect to* $\mathbf{W}$*.*

Exercise 13.3. Prove Prop. 13.2.

Note that, although an inverted matrix plays a role, the computation of harmonic Ritz pairs does not require matrix inversion.

Below, in Exercise 13.4, we analyse how Ritz values and harmonic Ritz values relate to their Ritz vectors and harmonic Ritz vectors, respectively.

Exercise 13.4. **Ritz values and harmonic Ritz values.** Let $\mathbf{A}$ be an $n \times n$ matrix and $\mathbf{u}$ a normalised n-vector. To ease notation, take $\tau = 0$.

(a) Show that $\mathbf{u}^*\mathbf{A}\mathbf{u} = \operatorname{argmin}_\vartheta \|\vartheta\mathbf{u} - \mathbf{A}\mathbf{u}\|_2$.

Let ϑ and $\widetilde{\vartheta}$ be such that $\mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathbf{u}$ and $\mathbf{A}\mathbf{u} - \widetilde{\vartheta}\mathbf{u} \perp \mathbf{A}\mathbf{u}$.

(b) Show that

$$\vartheta = \mathbf{u}^*\mathbf{A}\mathbf{u} \quad \text{and} \quad \widetilde{\vartheta} = \frac{\mathbf{u}^*\mathbf{A}^*\mathbf{A}\mathbf{u}}{\mathbf{u}^*\mathbf{A}^*\mathbf{u}}:$$

ϑ is a **Rayleigh quotient** associated with $\mathbf{u}$, and $\widetilde{\vartheta}$ is an **harmonic Rayleigh quotient** or **Temple quotient**. Note that Ritz values are Rayleigh quotients, while harmonic Ritz values are harmonic Rayleigh quotients.

(c) Conclude that

$$\frac{\vartheta}{\widetilde{\vartheta}} = \frac{|\mathbf{u}^*\mathbf{A}\mathbf{u}|^2}{\|\mathbf{A}\mathbf{u}\|_2^2} \quad \text{and} \quad \vartheta = r\widetilde{\vartheta} \quad \text{for some} \quad r \in [0,1], \tag{13.1}$$

$$|\vartheta| \leq \|\mathbf{A}\mathbf{u}\|_2 = \sqrt{|\vartheta|\,|\widetilde{\vartheta}|} \leq |\widetilde{\vartheta}| \quad \text{and} \quad \|\mathbf{A}\mathbf{u} - \vartheta\mathbf{u}\|_2 = \sqrt{\vartheta(\widetilde{\vartheta} - \vartheta)} \leq \tfrac{1}{2}|\widetilde{\vartheta}|, \tag{13.2}$$

$$|\vartheta| \leq \|\mathbf{A}\|_2 \quad \text{and} \quad |\widetilde{\vartheta}| \geq \frac{1}{\|\mathbf{A}^{-1}\|_2}. \tag{13.3}$$

Now, assume that $\mathbf{A}\mathbf{v}_j = \lambda_j \mathbf{v}_j$ with $(\mathbf{v}_j)$ orthonormal basis (or, equivalently, $\mathbf{A}$ is a normal matrix). Let

$$\mathbf{u} = \sum_j \beta_j \mathbf{v}_j \quad \text{and} \quad \mathbf{A}\mathbf{u} = \sum_j \gamma_j \mathbf{v}_j \quad \text{with} \quad \gamma_j \equiv \beta_j \lambda_j.$$

(d) Show that the Rayleigh quotient of $\mathbf{u}$ is a weighted average of the eigenvalues,

$$\vartheta = \frac{\sum \lambda_j |\beta_j|^2}{\sum_j |\beta_j|^2} = \sum_j \lambda_j \alpha_j \quad \text{with} \quad \alpha_j \equiv \frac{|\beta_j|^2}{\sum_i |\beta_i|^2}, \tag{13.4}$$

while the harmonic Rayleigh quotient is an *harmonic* weighted average of the eigenvalues (which explains the namings 'harmonic Rayleigh quotient' and 'harmonic Ritz value'),

$$\frac{1}{\widetilde{\vartheta}} = \frac{\sum_j \frac{1}{\lambda_j} |\gamma_j|^2}{\sum_j |\gamma_j|^2} = \sum_j \frac{1}{\lambda_j} \widetilde{\alpha}_j \quad \text{with} \quad \widetilde{\alpha}_j \equiv \frac{|\gamma_j|^2}{\sum_i |\gamma_i|^2}. \tag{13.5}$$

Assume $\mathbf{A}$ is normal and consider the situation of Exercise 13.4(d).

Note that ϑ can be close to and eigenvalue, say λ_1, while β_1 is very small (if, for instance, $\lambda_1 = 0$, $\lambda_2 = -1$, and $\lambda_3 = 1$, then $0 = \frac{1}{2}\lambda_2 + \frac{1}{2}\lambda_3$: $\mathbf{u} = \frac{1}{\sqrt{2}}(\mathbf{v}_2 + \mathbf{v}_3)$ is far away from $\mathbf{v}_1$ while the Ritz value $\vartheta = \mathbf{u}^*\mathbf{A}\mathbf{u} = 0$ is an excellent approximation of λ_1). If, however, λ_1 is an extremal eigenvalue, i.e., a vertex of the convex hull of the spectrum of $\mathbf{A}$, then, from (13.4) we learn that, a Ritz value ϑ can be close to λ_1 only if β_1 is relatively large, that is, only if the associated Ritz vector approximates the eigenvector $\mathbf{v}_1$ well. In other words, if we select extremal Ritz values (extremal in set of all Ritz values of order k) to approximate extremal eigenvalues, then we have a good approximation only if the Ritz vectors form good approximations of the eigenvectors. If we select an absolute small Ritz value to approximate an absolute small eigenvalues, then there is no guarantee that the Ritz vector forms a good approximation of the eigenvector with this absolute small eigenvalue. In such a case (13.5) tells us that it is safe to select for absolute small harmonic Ritz values. From (13.5) we learn that an absolute small harmonic Ritz value is close to an absolute small eigenvalue only of the harmonic Ritz vector is close to the eigenvector. The following proposition summarises these conclusions. Though we argued for normal matrices, the conclusion seems to be useful for general matrices as well.

Property 13.3 *Extremal Ritz pair can safely be selected if extremal eigenpairs are to be computed, harmonic Ritz pairs are more appropriate for computing eigenpairs with eigenvalues close to some target τ in the interior of the spectrum.*

Harmonic Ritz values can be viewed as extremal Ritz values for an inverted matrix in case the matrix is normal. This allows us to translate results for extremal Ritz values into results for harmonic Ritz values. Here, we give the "Cauchy interlace" type for Hermitian $\mathbf{A}$:

Theorem 13.4 *Let $\mathbf{A}$ be an $n \times n$ Hermitian matrix with λ_p the eigenvalue of $\mathbf{A}$ closest to τ. Assume λ_p simple and $\lambda_p \neq \tau$. Let $\mathbf{V}$ be an $n \times k$ orthonormal matrix spanning $\mathcal{V}$. Assume $\mathcal{V}$ does not contain an eigenvector of $\mathbf{A}$.*

Let $\widetilde{\vartheta}_1, \ldots$ be the harmonic Ritz values of $\mathbf{A}$ with respect to $\mathcal{V}$ and target τ, ordered such that $\widetilde{\vartheta}_{\ell+1} \leq \ldots \leq \widetilde{\vartheta}_k < \lambda_p < \widetilde{\vartheta}_1 \leq \widetilde{\vartheta}_2 \leq \ldots \leq \widetilde{\vartheta}_\ell$. We count according to multiplicity and allow harmonic Ritz values to have value ∞. Then,

(a) **[Lehmann]** *the interval $(\tau, \widetilde{\vartheta}_j]$ contains at least j eigenvalues of $\mathbf{A}$,*

(b) *the interval $(\tau, \widetilde{\vartheta}_j]$ contains at least j Ritz values of $\mathbf{A}$ w.r.t. $\mathcal{V}$.*

(c) *If $\mathcal{V}$ is a Krylov subspace, then Ritz values and harmonic Ritz values interlace.*

(d) *The inertia of all Ritz values equals the inertia of all $(1/\widetilde{\vartheta}_j)$.*[2]

[2]The inertia of a sequence (μ_j) of real μ_j is the triple $(\#\{j \,\big|\, \mu_j < 0\}, \#\{j \,\big|\, \mu_j = 0\}, \#\{j \,\big|\, \mu_j > 0\})$ as used in Sylvester's law of inertia, see Th. 12.3.

The intervals $(\tau, \widetilde{\vartheta}_j]$ are called **Lehmann intervals**. If only $\mathbf{V}$ and $\mathbf{AV}$ is available, then the Lehmann intervals are the smallest intervals with the property mentioned in (a) of the above theorem.

Exercise 13.5. Consider the situation of Th. 13.4.

(a) Harmonic Ritz values are allowed to have value ∞ (if $d \equiv \dim(\mathcal{V} \cup \mathcal{V}^\perp) > 0$) with multiplicity d. Argue that this definition of multiplicity for an harmonic Ritz value ∞ is consistent with the standard definition of multiplicity.

(b) Prove (d) of Th. 13.4. In particular, the number of zero Ritz values equals the number if ∞ harmonic Ritz values. (Hint: observe that $\mathbf{V}^*\mathbf{A}^*\mathbf{AV}$ allows a Cholesky decomposition, $\mathbf{V}^*\mathbf{A}^*\mathbf{AV} = LL^*$ with L $k \times k$ non-singular.)

If an (harmonic) Ritz value is close to an eigenvalue, then sometimes a few extra digits of accuracy can be obtained by the 'refined Ritz approach': if ϑ is an approximate eigenvalue, then

$$\mathbf{u} = \text{argmin}\{\|\mathbf{A}\widetilde{\mathbf{u}} - \tau\widetilde{\mathbf{u}}\|_2 \,\Big|\, \widetilde{\mathbf{u}} \in \mathcal{V}, \|\widetilde{\mathbf{u}}\|_2 = 1\}.$$

defines the **refined Ritz vector u**. Note that, given an approximate eigen*vector*, the Rayleigh quotient minimises the residual, while a refined Ritz vector minimises the residual for a given approximate eigen*value*.

The fact that harmonic Ritz values can be viewed as extremal Ritz values for an inverted matrix was use as an argument to favour harmonic Ritz values for selecting appropriate approximate interior eigenvalues. In the more specific Hermitian case, squaring can do same, it can turn interior eigenvalues to exterior ones, while preserving eigenvectors: if λ is an eigenvalue closest to τ (in the interior of the spectrum), then $(\lambda - \tau)^2$ is an extremal eigenvalue of $(\mathbf{A} - \tau\mathbf{I})^2$. It turns out that refined Ritz vectors can be viewed as Ritz vectors for the shifted squared matrix (see (13.8) in Exercise 13.6). In this Hermitian case, refined Ritz vectors and harmonic Ritz vectors are the same for carefully selected shifts (one for harmonic and one for refined Ritz; see (13.9)). This result can also be used to explain the favourable properties of harmonic Ritz values and to translate estimates on the accuracy of extremal Ritz vectors (as (12.12) and (12.13)) to estimates for harmonic Ritz vectors as is done in the next theorem.

Theorem 13.5 *Let* $\mathbf{A}$ *be an* $n \times n$ *Hermitian matrix with* λ_p *the eigenvalue of* $\mathbf{A}$ *closest to* τ. *Assume* λ_p *simple. Let* $\mathbf{V}$ *be an* $n \times k$ *orthonormal matrix spanning* $\mathcal{V}$. *Let* $(\vartheta, \mathbf{u})$ *be the harmonic Ritz pair with* ϑ *closest to* τ: $\mathbf{Au} - \vartheta\mathbf{u} \perp (\mathbf{A} - \tau\mathbf{I})\mathbf{V}$ *and* $|\vartheta - \tau|$ *smallest. Suppose* $\vartheta \in (\lambda_p, \lambda_{p+1})$. *Let* λ_q *the eigenvalue of* $\mathbf{A}$, $\lambda_q \neq \lambda_p$ *closest to* $\frac{1}{2}(\tau + \vartheta)$. *Then*

$$\tan^2(\mathbf{u}, \mathbf{x}_p) \leq \frac{\lambda_p - \tau}{\lambda_q - \tau} \frac{\vartheta - \lambda_p}{\lambda_q - \vartheta}. \tag{13.6}$$

With $t \equiv \tan\angle(\mathcal{V}, \mathbf{x}_p)$ *we have that*

$$0 < \vartheta - \lambda_p \leq t^2 \max_i \frac{(\lambda_i - \tau)(\lambda_i - \lambda_p)}{(\lambda_p - \tau) + t^2(\lambda_i - \tau)}. \tag{13.7}$$

Substitution of (13.7) in (13.6) estimates the angle ϕ between the harmonic Ritz vector $\mathbf{u}$ and the desired eigenvector $\mathbf{x}_p$ in terms of (the tangent t of) the angle ϕ_p between the search subspace $\mathcal{V}$ and $\mathbf{x}_p$: $\tan(\phi)$ is of order $t \equiv \tan(\phi_p)$ for t small provided that $\tau \neq \lambda_p$. Numerical results indicate that $\tan(\phi)$ is at most $\sqrt{k}$ times t. Apparently, even if the shift τ happens to be equal (or almost equal) to an eigenvalue, then one of the harmonic Ritz vectors forms an accurate approximation to the desired eigenvector. But in such a case selection based on harmonic Ritz values may have to be adapted. We will not further pursue this issue here.

Exercise 13.6. Let $\mathbf{A}$ and $\mathbf{V}$ be as in Theorem 13.5. To ease notation, take $\tau = 0$. For a scalar $\vartheta \in \mathbb{R}$ and vector $\mathbf{u} = \mathbf{V}y \in \mathbb{C}^n$, $\|\mathbf{u}\|_2 = 1$, consider the statements on Ritz, harmonic Ritz, and Refined Ritz pairs, respectively:

$$\begin{aligned} \mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathcal{V} \quad &\Leftrightarrow \quad \mathbf{V}^*\mathbf{A}\mathbf{V}y = \vartheta y \\ \mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathbf{A}\mathcal{V} \quad &\Leftrightarrow \quad \mathbf{V}^*\mathbf{A}^2\mathbf{V}y = \vartheta\mathbf{V}^*\mathbf{A}\mathbf{V}y \\ \mathbf{u} = \operatorname{argmin}\{\|\mathbf{A}\widetilde{\mathbf{u}}\|_2 \,\big|\, \widetilde{\mathbf{u}} \in \mathcal{V}, \|\widetilde{\mathbf{u}}\|_2 = 1\} \quad &\Leftrightarrow \quad \mathbf{V}^*\mathbf{A}^2\mathbf{V}y = \vartheta y. \end{aligned} \tag{13.8}$$

In the last statement, we assume ϑ to be the smallest eigenvalue of $\mathbf{V}^*\mathbf{A}^2\mathbf{V}$. Moreover, it then holds that $\vartheta = \min\{\|\mathbf{A}\widetilde{\mathbf{u}}\|_2^2 \,\big|\, \widetilde{\mathbf{u}} \in \mathcal{V}, \|\widetilde{\mathbf{u}}\|_2 = 1\}$.

(a) Prove the first two equivalencies.

(b) Prove the third equivalency and the characterisation of ϑ for the refined Ritz case (Hint: $\mathbf{u} = \operatorname{argmin}\ldots \;\Rightarrow\; \mathbf{A}^2\mathbf{u} \perp (\mathcal{V} \cap \mathbf{u}^\perp) \;\Rightarrow\; \mathbf{A}^2\mathbf{u} - \delta\mathbf{u} \perp \mathcal{V}$ for some $\delta \in \mathbb{R} \;\Rightarrow\; \delta = \|\mathbf{A}\mathbf{u}\|_2^2$).

(c) Prove that the Ritz vectors form a system that is orthogonal as well as A-orthogonal, the harmonic Ritz vectors are A-orthogonal and A^2-orthogonal, while the refined Ritz vectors (i.e., all vectors $\mathbf{V}y$ that solve $\mathbf{V}^*\mathbf{A}^2\mathbf{V}y = \vartheta\, y$ for some $\vartheta \in \mathbb{R}$) are orthogonal and A^2-orthogonal.

(d) Put $\sigma \equiv \vartheta/2$. Prove the following equivalency

$$\mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathbf{A}\mathcal{V} \quad \Leftrightarrow \quad \mathbf{u} = \operatorname{argmin}\{\|\mathbf{A}\widetilde{\mathbf{u}} - \sigma\widetilde{\mathbf{u}}\|_2 \,\big|\, \widetilde{\mathbf{u}} \in \mathcal{V}, \|\widetilde{\mathbf{u}}\|_2 = 1\}. \tag{13.9}$$

Switching to refined Ritz vectors can be dangerous if the approximate eigenvalue is not close to the wanted eigenvalue: the refined Ritz vector may be close to some eigenvector of which the eigenvalue happens to be (accidentally) close to the approximate eigenvalue and may have no component of the wanted eigenvector.

Refined Ritz vectors are singular vectors with smallest singular value (see Exercise 13.7(c)).

The following exercise shows how Ritz pairs and harmonic Ritz pairs can be computed for general subspaces $\mathcal{V}$ (= span($\mathbf{V}$)). The approach may lead to a **generalised eigenvalue problem**, that is, a problem of the form $Ax = \lambda Bx$ with A and B $k \times k$ matrices. If, say B, is non-singular, then this generalised problem is equivalent to a 'standard' eigenvalue problem, $B^{-1}Ax = \lambda x$. However, it is not always stable to recast the generalised problem into standard form. The so-called **QZ-algorithm** brings the generalised problem (of low dimension) into the form $Ty = \lambda Sy$ with T and S upper triangular matrices, using unitary matrices: $T = Z^*AQ$ and $S = Z^*BQ$ with Z and Q $k \times k$ unitary. Similar to the QR-algorithm for $Ax = \lambda x$, the S and T appear in the QZ-algorithm as limit of repeated basis rotations in image space (resulting in a basis Z) and in domain space (resulting in a basis Q) applied to A and B, respectively.

Exercise 13.7. Let $\mathbf{V}$ be $n \times k$ orthogonal with columns spanning the search subspace $\mathcal{V}$. Let $\mathbf{A}\mathbf{V} = \mathbf{W}R$ be the QR-decomposition of $\mathbf{A}\mathbf{V}$, i.e., $\mathbf{W}$ is $n \times k$ orthogonal and R $k \times k$ upper triangular. Put $M \equiv \mathbf{V}^*\mathbf{W}$. Below $\mathbf{u} = \mathbf{V}y$.

(a) Show that $\mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathbf{V} \quad \Leftrightarrow \quad MRy = \vartheta y \quad$ ($(\vartheta, \mathbf{u})$ is a Ritz pair).
Note that the H of ALG. 13.1 equals MR.

(b) Show that $\mathbf{A}\mathbf{u} - \vartheta\mathbf{u} \perp \mathbf{A}\mathbf{V} \quad \Leftrightarrow \quad Ry = \vartheta M^*y \quad$ ($(\vartheta, \mathbf{u})$ is an harmonic Ritz pair).

(c) Show that $\mathbf{u} = \operatorname{argmin}\{\|\mathbf{A}\widetilde{\mathbf{u}} - \tau\widetilde{\mathbf{u}}\|_2 \,\big|\, \widetilde{\mathbf{u}} \in \mathcal{V}\} \quad \Leftrightarrow \quad y$ is the right singular vector of $R - \tau M^*$ associated to the smallest singular value of $R - \tau M$ ($\mathbf{u}$ is a refined Ritz vector).
Usually τ is selected to be a Ritz value.

B Sorting the Schur form

If H is the matrix $\mathbf{A}$ projected onto a low, k-dimensional subspace, then, for stability reasons, we will compute a Schur decomposition $H = USU^*$, with U unitary and S upper triangular, (cf., Th. 0.6 and Exercise 0.17), rather than an eigenvector decomposition $H = X\Lambda X^{-1}$ with Λ diagonal. To facilitate selection of interesting approximate eigenpairs, we may want to reorder the Schur decomposition. To be more precise, we want a Schur decomposition with an upper

triangular S with diagonal entries in specified order. Because, if $S_{11}, \ldots, S_{\ell\ell}$ are the eigenvalues (Ritz values) of interest, then the first ℓ columns of U form an orthonormal basis of the space spanned by the eigenvectors associated with these eigenvalues, while, say the kth column of U, does not span the eigenspace with eigenvalue S_{kk}. Reordering the eigenvector decomposition is easy: if P is a reordering of $(1, 2, \ldots, k)$, then $HX(:, P) = X(:, P)\Lambda(P, P)$ is an eigenvector decomposition with eigenvalues reorder on the diagonal of Λ. Reordering the Schur form requires more care: $S(P, P)$ is in general not upper triangular. The following property states that reordering is possible even with unitary matrices, which is in line with stability demands. The following exercise describes how this can be done in practice.

Property 13.6 *For a $k \times k$ upper triangular matrix S and a permutation π of $(1, \ldots, k)$ (i.e., π is a bijection onto $\{1, 2, \ldots, k\}$), there is a $k \times k$ unitary matrix Q such that $S' \equiv Q^* S Q$ upper triangular and the ith diagonal entry S'_{ii} equals $S_{\pi(i)\pi(i)}$ $(i = 1, \ldots, k)$.*

Exercise 13.8. **Sorting the Schur decomposition.** Let S and π be as in Prop. 13.6.

(a) Consider the (complex) 2×2 matrix

$$\begin{bmatrix} \lambda & \alpha \\ 0 & \mu \end{bmatrix}.$$

Show that there is a Givens rotation (cf., Exercise 3.5)

$$\begin{bmatrix} c & s \\ -\overline{s} & c \end{bmatrix} \quad \text{with } c \in [0, 1] \text{ en } s \in \mathbb{C} \text{ such that } \; c^2 + |s|^2 = 1$$

($c = \cos(\phi)$ and $s = \sin(\phi)$) such that

$$\begin{bmatrix} c & s \\ -\overline{s} & c \end{bmatrix} \begin{bmatrix} \lambda & \alpha \\ 0 & \mu \end{bmatrix} \begin{bmatrix} c & -s \\ \overline{s} & c \end{bmatrix} = \begin{bmatrix} \mu & \beta \\ 0 & \lambda \end{bmatrix}$$

for some appropriate scalar β. With $t \equiv s/c$ we should have that $t = (\mu - \lambda)/\alpha$. Hence,

$$t = (\mu - \lambda)/\alpha, \quad c = 1/\sqrt{|t|^2 + 1}, \quad \text{and} \quad s = ct.$$

(b) Consider S. Suppose we want to 'switch' the 1th and the kth diagonal element: we are interested in a unitary matrix Q such that $S' \equiv Q^* S Q$ is upper triangular such that $S'_{ii} = S_{ii}$ if $i \notin \{1, k\}$ and $S'_{11} = S_{kk}$ and $S'_{kk} = S_{11}$. Show that this can be achieve with Givens rotations. How many Givens rotations are required?

Any permutation can be expressed as a product of basic permutation, where a basic permutation is a permutation that switches two indices only.

(c) Show that there is a unitary matrix Q such that with $S' \equiv Q^* S Q$, S' is upper triangular, and the (i, i) entry of S' equals the $(\pi(i), \pi(i))$ entry of S. Q can be obtained as a product of Givens rotations.

(d) Write a Matlab function subroutine

```
[S1,Q1]=SortS(S,p)
```

that sorts the upper triangular matrix S according to the permutation π: the matrices are $k \times k$, S is upper triangular, the matrices $S_1 \equiv$`S1` and $Q_1 \equiv$`Q1` are such that S_1 is upper triangular, Q_1 is unitary, $S = Q_1 S_1 Q_1^*$ and with `D=diag(S)`, `D1=diag(S1)` we have that `D1=D(p)`, where `p` is a reordering of `1:k` according to π.

C Thick restart

Consider a subspace method for computing an eigenpair. Let k be the dimension of the search subspace $\mathcal{V}_k$ at a certain point in the computational process. If k is large (typically $k = k_{\max} = 60$), then the demands for storing a basis $\mathbf{v}_1, \ldots, \mathbf{v}_k$ of the search subspace and the costs for computing a well conditioned expansion vector $\mathbf{v}_{k+1}$ (i.e., the costs of orthogonalising in some sense the expansion vector to $\mathbf{v}_1, \ldots, \mathbf{v}_k$) will be high and we may want to continue the computational process with a smaller search subspace $\widetilde{\mathcal{V}}_\ell$, that is, with a subspace of dimension ℓ with $\ell < k$. An obvious choise for such a smaller subspace is the one spanned by the ℓ 'best' eigenvector approximations that can be extracted from the larger subspace. If Ritz pairs $(\vartheta_j, \mathbf{u}_j)$ are selected for approximating eigenpairs, then one may select the ℓ Ritz vectors, say $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$, with Ritz value that best have a desirable property. For instance, if one is interested in the eigenvalue with largest real part or the one closest to some target value τ (in which case one will select harmonic Ritz pairs rather than Ritz pairs, cf., Lecture 13A), then the ℓ Ritz values will be the ones with largest real part ($\mathrm{Re}(\vartheta_j) \geq \mathrm{Re}(\vartheta_{j+1})$) or the ones closest to τ ($|\vartheta_j - \tau| \leq |\vartheta_{j+1} - \tau|$), respectively. With $\ell = 1$ we have the standard restart approach. However, **thick restart**, that is, with a larger ℓ, usually is more effective (typically $\ell = k_{\min} = 15$): the error in the Ritz vector $\mathbf{u}_1$ (as eigenvector approximation) usually has largest component in $\mathbf{u}_2, \ldots$. Therefore, discarding $(\vartheta_2, \mathbf{u}_2)$ at restart will slow down convergence towards the first wanted eigenpair. Moreover, in eigenvalue computation, one is often interested in computing more than one eigenpair. The Ritz pair $(\vartheta_2, \mathbf{u}_2)$ is typically converging towards the second wanted eigenpair. Therefore, maintaining this Ritz vector in the search subspace will provide a good starting vector later on in the computational process when the first eigenpair has been detected and the search for the next eigenpair is started.

A thick restart procedure for the Arnoldi process requires additional care. We not only want a subspace spanned by the appropriate approximate eigenvectors, but we also need a basis of this subspace that forms an Arnoldi relation.

Assume that, for $k = k_{\max}$,

$$\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k, \quad \mathbf{v}_1 \equiv \mathbf{V}_k e_1 \tag{13.10}$$

is an Arnoldi relation: $\mathbf{V}_j$ is $n \times j$ orthonormal, $\underline{H}_k$ is $(k+1) \times k$ upper Hessenberg. With H_k the $k \times k$ upper block of $\underline{H}_k$, let

$$H_k = Q_k S_k Q_k^*$$

be the Schur decomposition of H_k, that is, S_k is $k \times k$ upper triangular and Q_k is $k \times k$ unitary (see Th. 0.6). In particular, as before, the index k of $\mathbf{V}_k$, $\underline{H}_k$, ... refers to the number of columns of these matrices. Let the Schur decomposition be sorted such that $\vartheta_1, \ldots, \vartheta_\ell, \vartheta_{\ell+1}, \ldots, \vartheta_k$ be the on the diagonal of S_k (see Prop. 13.6 and Exercise 13.8): the first ℓ diagonal entries of S are the Ritz values of interest. Let $\mathbf{u}_i$ be the Ritz vectors associated with the Ritz values ϑ_i. Let $\widetilde{\mathbf{V}}_\ell$ the subspace spanned by $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$.

At restart, an Arnoldi relation $\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell$ is required: we need an orthonormal basis of $\widetilde{\mathcal{V}}_\ell$ that forms the columns of $\widetilde{\mathbf{V}}_\ell$ and for which $\underline{\widetilde{H}}_\ell$ is upper Hessenberg. In particular, $\widetilde{\mathbf{V}}_\ell$ spans a Krylov subspace. The following property states that such a basis exists. For a construction (and a proof of the property), we refer to the following exercise.

Theorem 13.7 *Any set $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ of Ritz vectors in a Krylov subspace $\mathcal{K}_k(\mathbf{A}, \mathbf{v}_1)$ spans a Krylov subspace $\mathcal{K}_\ell(\mathbf{A}, \widetilde{\mathbf{v}}_1)$ of $\mathbf{A}$ (for some n-vector $\widetilde{\mathbf{v}}_1$).*

Exercise 13.9. Restarting Arnoldi method. Consider the situation as described above. In particular, $\widetilde{\mathcal{V}}_\ell$ is the space spanned by the ℓ Ritz vectors $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$ associated to the first ℓ Ritz values.

(a) Prove that $\mathbf{A}\mathbf{u}_i - \vartheta_i \mathbf{u}_i$ is a multiple of $\mathbf{v}_{k+1}$.

(b) Let $\mathbf{W}_\ell$ be the matrix consisting of the first ℓ columns of $\mathbf{V}_k\, Q_k$. Show that the columns of $\mathbf{W}_\ell$ form an orthonormal basis if $\mathcal{V}_\ell$.

(c) Show that

$$\mathbf{A}\mathbf{W}_\ell = [\mathbf{W}_\ell, \mathbf{v}_{k+1}]\, \underline{G}_\ell \quad \text{for some } (\ell+1)\times\ell \text{ matrix } \quad \underline{G}_\ell = \begin{bmatrix} S_\ell \\ c^* \end{bmatrix} \tag{13.11}$$

for some ℓ-vector c. Here, S_ℓ is the $\ell \times \ell$ left upper block of S_k. Describe c.

(d) Write a MATLAB function subroutine (cf., the procedure in Exercise 5.27, Exercise 5.27(d))

```
[tQ,tH]=Hessenberg(G)
```

that for a given $(\ell+1)\times\ell$ matrix $\underline{G}_\ell$ (`G=`$\underline{G}_\ell$) produces an $(\ell+1)\times\ell$ upper Hessenberg matrix $\underline{\widetilde{H}}_\ell$ (`tH=`$\underline{\widetilde{H}}_\ell$) with an $\ell\times\ell$ unitary matrix Q (`tQ=`Q with Q a product of Householder reflections) and such that $\underline{\widetilde{H}}_\ell\, Q = Q^\circ\, \underline{G}_\ell$. Here Q° is the matrix Q extended with a row of zeros and the column $e_{\ell+1}$.

(e) Turn $\mathbf{A}\mathbf{W}_\ell = [\mathbf{W}_\ell, \mathbf{v}_{k+1}]\, \underline{G}_\ell$ into an Arnoldi relation

$$\mathbf{A}\widetilde{\mathbf{V}}_\ell = [\widetilde{\mathbf{V}}_\ell\,, \widetilde{\mathbf{v}}_{\ell+1}]\, \underline{\widetilde{H}}_\ell,$$

i.e., the columns of $\widetilde{\mathbf{V}}_\ell$ form an orthonormal Krylov basis for $\widetilde{\mathcal{V}}_\ell$. What can you tell about $\widetilde{\mathbf{v}}_{\ell+1}$?

With Arnoldi relation $\mathbf{A}\mathbf{V}_k = [\mathbf{V}_k, \mathbf{v}_{k+1}]\, \underline{H}_k$ of order k, restart requires the following steps

1) Compute the Schur decomposition $H_k Q = QS$ of the $k\times k$ upper block H_k of $\underline{H}_k$

```
K=1:k; [Q,S]=schur(H(K,:));
```

2) Sort S: $SQ_1 = Q_1 S_1$ with diagonal S_1 in prescribed order (cf., Exercise 13.8)

```
[S1,Q1]=SortS(S,p);
```

3) Compute an orthonormal basis $\mathbf{W}$ of the span of the first ℓ Ritz vectors

```
L=1:ell; W=V*Q*Q1(:,L);
```

4) Turn the left $\ell\times\ell$ upper block of S_1 extended with the proper part of the last row of $\underline{H}_k$ (properly represented) into an upper Hessenberg matrix $\underline{\widetilde{H}}_\ell$ (cf., Exercise 13.9)

```
c=H(k+1,:)*Q*Q1(:,L); G=[S1(L,L);c]; [tQ,tH]=Hessenberg(G);
```

5) Compute an orthonormal Krylov basis $\widetilde{\mathbf{V}}$ of span($\mathbf{W}$)

```
tV=W*tQ;
```

Then $\mathbf{A}\widetilde{\mathbf{V}}_\ell = [\widetilde{\mathbf{V}}_\ell, \mathbf{v}_{k+1}]\, \underline{\widetilde{H}}_\ell$ is an Arnoldi relation of order ℓ and $\widetilde{\mathbf{V}}_\ell$ spans the space spanned by the first ℓ Ritz vectors of $\mathbf{A}$ with respect to span($\mathbf{V}_k$).

Note that 3) and 5) can be combined: `V=V*(Q*Q1(:,L)*tQ)`, thus limiting the number of high dimensional operations. This approach is called the **Arnoldi-Schur restart**.

The above thick restart strategy, restart with a space spanned by Ritz vectors of interest, is essentially applicable to any subspace method (not necessarily a Krylov subspace method as Lanczos and Arnoldi). If no Arnoldi decomposition is required (as in Jacobi–Davidson methods) then, at restart, step 4) and 5) can be skipped. For such methods, H_k will be the matrix of $\mathbf{A}$ projected onto the search subspace. In general H_k will not be Hessenberg.

Implicitly Restarted Arnoldi Method (IRAM)

The well-know **implicitly restarted Arnoldi method** (**IRAM**) is an elegant and efficient implementation of the above steps. It combines the steps 2) and 4) (and 3) and 5)) (sort, select and form Krylov basis for the selected Ritz vectors). It exploits the fact that the QR-algorithm for computing the Schur decomposition for dense matrices actually incorporates a polynomial filter in the first column. In IRAM, the steps of the QR-algorithm are applied to the Hessenberg matrix $\underline{H}_k$. This matrix is not square and as a consequence, each QR-RQ step reduces the size by one. After $k-\ell$ of such steps, the Arnoldi decomposition of order k is reduced to an Arnoldi decomposition of order ℓ. The first column of the reduced Arnoldi matrix $\widetilde{\mathbf{V}}_\ell$ equals $p(\mathbf{A})\mathbf{v}_1$,

Let $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$ be an Arnoldi relation
Set $U = I_{k+1}$, $\underline{S} = \underline{H}_k$.
For $j = k, k-1, \ldots, \ell+1$ do
  1) Select a shift μ_j.
  2) Compute the QR-decomposition $\underline{S} - \mu_j \underline{I}_j = \underline{Q}R$.
  3) Compute $\underline{S} \leftarrow R\underline{Q}' + \mu_j \underline{I}_{j-1}$
  4) $U \leftarrow U\underline{Q}$.
end for
$\widetilde{\mathbf{V}}_{\ell+1} = \mathbf{V}_{k+1}U$, $\underline{\widetilde{H}}_\ell = \underline{S}$

ALGORITHM 13.2. **Implicit polynomial filtering of Arnoldi relations**. A given Arnoldi relation $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$ for a general square matrix $\mathbf{A}$ is reduced to an Arnoldi relation $\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell$ such that the first column $\widetilde{\mathbf{v}}_1$ of $\widetilde{\mathbf{V}}_\ell$ equals a multiple of $p(\mathbf{A})\mathbf{v}_1$ where $\mathbf{v}_1$ is the first column of $\mathbf{V}_k$ and p is the polynomial $p(\lambda) = (\lambda - \mu_{\ell+1}) \cdot \ldots \cdot (\lambda - \mu_k)$ as defined by the μ_j. See Theorem 13.8.
Matrices denotes as $\underline{B}$ are of size $(j+1) \times j$, i.e., square plus one additional row. If $\underline{B}$ is $(j+1) \times j$, then $\underline{B}'$ denotes the $j \times (j-1)$ left upper block of $\underline{B}$: $\underline{B}'$ is the matrix that arises by removing the last row and last column of $\underline{B}$. $\underline{I}_{j-1}$ is the left $j \times j-1$ block of the $j \times j$ identity matrix. Note that the index j in the 'For'-loop decreases.

where $\mathbf{v}_1$ is the first column of the Arnoldi matrix $\mathbf{V}_k$ of order k, and p is the polynomial as defined by the (shifted) QR-RQ steps.

In ALG. 13.2, a limited number of steps $(k - \ell)$ of the shifted QR-algorithm is applied to the matrix $\underline{H}_k$. This has the same effect as applying a polynomial filter (of degree m) in $\mathbf{A}$ to $\mathbf{v}_1$: the filtered vector $p(\mathbf{A})\mathbf{v}_1$ rather than $\mathbf{v}_1$ shows up as the initial vector in an Arnoldi relation. Here, p is a polynomial of degree m. This result is formulated in Theorem 13.8. For a proof, we refer to Exercise 13.10. The filtering is *implicit* in the sense that it is obtained by applying the filter to $\underline{H}_k$, i.e., in low dimensional space, rather than to $\mathbf{A}$, i.e., in high dimensional space. Vector updates (at the end of ALG. 13.2) are the only high dimensional operations that are required for this, no MVs, no inner products. In Theorem 13.9, we will see that this implicit polynomial filtering approach will provide us with the elegant way as indicated above for restarting Arnoldi with a space of selected Ritz vectors.

Theorem 13.8 *Consider* ALG. 13.2. *With*

$$p(\zeta) \equiv (\zeta - \mu_{\ell+1}) \cdot \ldots \cdot (\zeta - \mu_{k-1})(\zeta - \mu_k) \qquad (\zeta \in \mathbb{C}),$$

$\mathbf{v}_1 \equiv \mathbf{V}_k e_1$, *and* $\widetilde{\mathbf{v}}_1 \equiv \widetilde{\mathbf{V}}_\ell e_1$, *we have*

$$\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell \quad \textit{and} \quad \widetilde{\mathbf{v}}_1 = \widetilde{\tau}\, p(\mathbf{A})\mathbf{v}_1 \quad \textit{for some } \widetilde{\tau} \in \mathbb{C}.$$

Exercise 13.10. ***Proof of Theorem*** 13.8**.** Consider ALG. 13.2.

Select a shift $\mu \in \mathbb{C}$. Form the QR-decomposition of $\underline{H}_k - \mu \underline{I}_k$:

$$\underline{H}_k - \mu \underline{I}_k = \underline{Q}_k R_k,$$

where $\underline{Q}_k$ is $(k+1) \times k$ orthonormal and R_k is $k \times k$ upper triangular.

(a) Show that such a decomposition exists. Show that $\underline{Q}_k$ is upper Hessenberg.

Let $\underline{Q}_k{}'$ be the $k \times (k-1)$ left upper block of $\underline{Q}_k$. Define

$$\underline{H}^+_{k-1} \equiv R_k \underline{Q}_k{}' + \mu \underline{I}_{k-1} \quad \text{and} \quad \mathbf{V}^+_{k-1} \equiv \mathbf{V}_k \underline{Q}_k{}', \quad \mathbf{V}^+_k \equiv \mathbf{V}_{k+1}\underline{Q}_k.$$

(b) Prove that $\underline{H}^+_{k-1}$ is $k \times (k-1)$ upper Hessenberg. Prove that

$$\underline{Q}_k \, \underline{H}^+_{k-1} = \underline{H}_k \, \underline{Q}_k{}' \quad \text{and} \quad \mathbf{A}\mathbf{V}^+_{k-1} = \mathbf{V}^+_k \underline{H}^+_{k-1}, \quad (\mathbf{A} - \mu\,\mathbf{I})\mathbf{V}_k = \mathbf{V}^+_k R_k.$$

(c) Show that

$$(\mathbf{A} - \mu\,\mathbf{I})\mathbf{v}_1 = \tau\,\mathbf{v}^+_1 \qquad \text{for some } \tau \in \mathbb{C}.$$

Here, $\mathbf{v}_1$ and $\mathbf{v}^+_1$ are the first columns of the matrix $\mathbf{V}_k$ and $\mathbf{V}^+_k$, respectively. Conclude that the columns of $\mathbf{V}^+_k$ form an orthonormal Krylov basis for the Krylov subspace of order k generated by $\mathbf{A}$ and the vector $(\mathbf{A} - \mu\,\mathbf{I})\mathbf{v}_1$.

(d) Let $U \equiv \underline{Q}_k \underline{Q}_{k-1} \ldots \underline{Q}_{\ell+1}$ with $\underline{Q}_j = \underline{Q}$ of step j ($j = k, k-1, \ldots, \ell+1$) in ALG. 13.2. Show that U is $(k+1) \times (\ell+1)$ orthonormal, $U' = \underline{Q}'_k \underline{Q}'_{k-1} \ldots \underline{Q}'_{\ell+1}$, U' is $k \times \ell$ orthonormal, and $U\underline{\widetilde{H}}_\ell = \underline{H}_k U'$.

(e) Prove Theorem 13.8.

Implicit polynomial filtering of a Arnoldi relation is useful in itself. With a specific choice of shifts, we have the restart strategy of IRAM: select 'unwanted' Ritz values for the shifts μ_j. The polynomial filter method removes the selected Ritz values from the Arnoldi relation, i.e., the eigenvalues of the 'new' Hessenberg matrix $\widetilde{H}_\ell$ are precisely the 'wanted' Ritz values $\vartheta_1, \ldots, \mu_\ell$, that is, the non-selected eigenvalues of H_k (as stated in the following theorem. As in ALG. 13.2, the index j in the theorem decreases).

Theorem 13.9 *Consider* ALG. 13.2. *Let $(\vartheta_j, \mathbf{u}_j)$ be the Ritz pairs for $\mathbf{A}$ from H_k and $\mathbf{V}_k$. If the selected shifts μ_j are equal to the Ritz values ϑ_j,*

$$\mu_j = \vartheta_j \qquad (j = k, k-1, \ldots, \ell+1),$$

then $\vartheta_1, \ldots, \vartheta_\ell$ are Ritz values of the Arnoldi relation $\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell$ and $\widetilde{\mathbf{V}}_\ell$ spans the space spanned by the Ritz vectors $\mathbf{u}_1, \ldots, \mathbf{u}_\ell$.

Exercise 13.11. ***Proof of Theorem*** 13.9**.** We continue Exercise 13.10.
Let $\vartheta_1, \ldots, \vartheta_k$ be the eigenvalues of H_k. Select $\mu_j = \vartheta_j$ for $j = k, k-1, \ldots, \ell+1$.

Recall that (cf., Th. 5.9, Th. 6.1, and Exercise 6.6) for a polynomial q of exact degree k,

$$q(H_k) = 0 \quad \Leftrightarrow \quad q(H_k)e_1 = 0 \quad \Leftrightarrow \quad q(\mathbf{A})\mathbf{v}_1 \perp \mathbf{V}_k \quad \Leftrightarrow \quad q(\vartheta_j) = 0 \quad (j = 1, \ldots, k).$$

Let q be the characteristic polynomial of H_k, i.e., $q(H_k) = 0$.

(a) For $j = 1, \ldots, k$, let q_j be the polynomial of degree $k-1$ such that

$$q(\zeta) = (\zeta - \vartheta_j)q_j(\zeta) \qquad (\zeta \in \mathbb{C}).$$

Prove that $\mathbf{w}_j \equiv q_j(\mathbf{A})\mathbf{v}_1$ is a multiple of the Ritz vector $\mathbf{u}_j$. (Hint: $(\mathbf{A} - \vartheta_j\mathbf{I})\mathbf{w}_j \perp \mathbf{V}_k$.)

(b) With $p(\zeta) \equiv (\zeta - \vartheta_{\ell+1}) \cdot \ldots \cdot (\zeta - \vartheta_k) \quad (\zeta \in \mathbb{C})$, there is a polynomial r of exact degree ℓ such that $q(\zeta) = r(\zeta)p(\zeta) \quad (\zeta \in \mathbb{C})$. Prove that $r(\mathbf{A})\widetilde{\mathbf{v}}_1 \perp \widetilde{\mathbf{V}}_\ell$.

(c) Prove Theorem 13.9.

Exercise 13.12. **Restart with Ritz vectors.** This exercise suggests an alternative proof of Th. 13.9 and gives some additional insight in polynomial filtering with Ritz values.

Let $\vartheta_1, \ldots, \vartheta_k$ be the eigenvalues of H_k.
We continue Exercise 13.10. Apply the steps in (a) and (b) of Exercise 13.10 with $\mu = \vartheta_k$.

(a) Show that $\underline{Q}_k\, e_k = e_{k+1}$.[3] Conclude that $\mathbf{V}^+_k e_k = \mathbf{v}_{k+1} \equiv \mathbf{V}_{k+1}\, e_{k+1}$ and $\widetilde{\mathbf{v}}_{\ell+1} \equiv \widetilde{\mathbf{V}}_{\ell+1}e_{\ell+1} = \mathbf{v}_{k+1}$ with $\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell$ the Arnoldi relation of Theorem 13.9 and Exercise 13.11.

[3] When MATLAB's `qr`, `[uQ,R]=qr(uH-μ*uI,0);`, has been used to compute $\underline{Q}_k$, cf., Exercise 13.10, then the right bottom element of $\underline{Q}_k$ may differ from 1 by a sign, i.e., by $e^{i\phi}$.

(b) Show that the eigenvalues of the $(k-1)\times(k-1)$ upper block H_{k-1}^{+} of $\underline{H}_{k-1}^{+}$ are equal to $\vartheta_1,\ldots,\vartheta_{k-1}$.

(c) As in (d) of Exercise 13.10, the $\underline{Q}_j$ can be assembled into a matrix U, $U = \underline{Q}_k \cdot \ldots \cdot \underline{Q}_{\ell+1}$, before updating $\mathbf{V}_{k+1}$. Show that the last row and last column of U consists of all 0s except for the $(k+1,\ell+1)$-entry which equals 1.[3]

Exercise 13.13. **Restart with harmonic Ritz vectors.** Let $\vartheta_1,\ldots,\vartheta_k$ be the harmonic Ritz values with respect to $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$. As with Ritz values as shifts, ALG. 13.2 with harmonic Ritz values leads to the Arnoldi relation $\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell$ with $\widetilde{\mathbf{V}}_\ell$ spanning the space spanned by the harmonic Ritz vectors that are not selected as shifts. This result can be proved by adapting the arguments of Exercise 13.11. Below we present an elementary proof of the fact that the residuals (or, more precise, a multiple of the residuals) of harmonic Ritz pairs (and of GMRES) is preserved in the restart. This result can be exploited in an implicit restart version of GMRES (not discussed here).

Let y_j be the harmonic pre-Ritz vector with harmonic Ritz value ϑ_j: $(\underline{H}_k - \vartheta_j \underline{I}_k)y_k \perp \underline{H}_k$ (see Exercise 6.4). Note that $\vartheta_j \neq 0$.
We continue Exercise 13.10. Apply the steps in (a) and (b) of Exercise 13.10 with $\mu = \vartheta_k$.

Let $\vec{\gamma}_{k+1}$ be the $(k+1)$-vector with first coordinate equal to 1 such that $\vec{\gamma}_{k+1}^{*}\underline{H}_k = \vec{0}_k^{*}$ (as in (6.2)). Recall that $(\underline{H}_k - \vartheta_j \underline{I}_k)y_k = \tau\vec{\gamma}_{k+1}$ for some scalar τ (see Exercise 6.4(f)).

(a) Let $\widetilde{\gamma}$ be a $(k+1)$-vector such that $\widetilde{\gamma}^{*}(\underline{H}_k - \vartheta_k \underline{I}_k) = \vec{0}^{*}$. Prove that $\widetilde{\gamma} \perp \vec{\gamma}_{k+1}$.

(b) Let $\underline{H}_k - \vartheta_k \underline{I}_k = \underline{Q}_k R_k$ be the QR-decomposition. Show that $\vec{\gamma}_{k+1} \in \mathcal{R}(\underline{Q}_k)$.

(c) Apply ALG. 13.2 with shifts $\mu_j = \vartheta_j$ harmonic Ritz values. Let $U \equiv \underline{Q}_k \underline{Q}_{k-1} \ldots \underline{Q}_{\ell+1}$ with $\underline{Q}_j$ the $\underline{Q}$ of step j ($j = k, k-1, \ldots, \ell+1$) in ALG. 13.2. Prove that $\vec{\gamma}_{k+1} \in \mathcal{R}(U)$.

Let $\mathbf{A}\widetilde{\mathbf{V}}_\ell = \widetilde{\mathbf{V}}_{\ell+1}\underline{\widetilde{H}}_\ell$ the Arnoldi relation resulting from ALG. 13.2 with harmonic Ritz values as shifts. Put $\widetilde{\gamma}_{\ell+1} \equiv U^{*}\vec{\gamma}_{k+1}$.

(d) Show that $\widetilde{\gamma}_{\ell+1}$ is an $(\ell+1)$-vector, $\widetilde{\gamma}_{\ell+1}^{*}\underline{\widetilde{H}}_\ell = 0^{*}$, $U\widetilde{\gamma}_{\ell+1} = \vec{\gamma}_{k+1}$. In particular $\widetilde{\mathbf{V}}_{\ell+1}\widetilde{\gamma}_{\ell+1} = \mathbf{V}_{k+1}\vec{\gamma}_{k+1}$.

In practice (when modified Gram–Schmidt has been used for orthonormalisation of the Arnoldi vectors), the columns of $\widetilde{\mathbf{V}}_\ell$ may loose orthogonality: modified Gram-Schmidt leads to loss of orthogonality exactly when with GMRES the size of the GMRES residual reaches machine precision. With implicit restarts orthogononality may be lost before the residual is small enough. This can be circumvented with a more stable form of Gram–Schnidt (as repeated Gram–Schmidt), or by reorthogonalizing the Arnoldi relation, as explained in the following exercise.

Exercise 13.14. Suppose $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$ is a Krylov flag, i.e., $\underline{H}_{k+1}$ is upper Hessenberg, the first k columns of $\mathbf{V}_{k+1}$ are the columns of $\mathbf{V}_k$.

(a) Let $\mathbf{V}_{k+1} = \mathbf{W}_{k+1}R_{k+1}$ be the QR-decomposition (in economical form) of $\mathbf{V}_{k+1}$. Show that $\mathbf{A}\mathbf{W}_k = \mathbf{W}_{k+1}R_{k+1}\underline{H}_k R_k^{-1}$ is an Arnoldi relation. Here, R_k is the $k\times k$ left upper block of R_{k+1}. Note that the first j columns of $\mathbf{V}_{k+1}$ span the same space as the first j columns of $\mathbf{W}_{k+1}$.

Exercise 13.15. **Implicitly restarted Lanczos (IRL).**
Consider the situation of Exercise 13.10. Assume that $\underline{H}_k$ is tridiagonal.

(a) Prove that $\underline{H}_{k-1}^{+}$ and $\underline{\widetilde{H}}_\ell$ are tridiagonal.

(b) Explain how a thick implicitly restart strategy can be incorporated in Lanczos.

(c) The strategy as indicated in (b) assumes storage of the Lanczos vectors. One of the advantages of Lanczos compared to Arnoldi is the low storage. This advantage is lost in implicitly restarted Lanczos. What are the advantages (if any) of IRL compared to IRAM?

D Deflation

If an eigenpair, say $(\lambda, \mathbf{x})$, has been detected (with a certain accuracy), the search will be continued to a next eigenpair. In order to avoid that the already detected eigenpair(s) will re-enter the computation (due to errors, rounding errors as well as approximation errors), these eigenpairs are to be deflated. Eigenvalues that re-enter the computation due to erros are called **ghost eigenvalues** or **spurious eigenvalues**. As we learnt from the power method, it is not sufficiently stable to apply deflation in one step only. We therefore deflate the problem (the matrix) rather than the initial vector.

For stability reasons, usually a partial Schur decomposition is computed, $\mathbf{A}\mathbf{Q}_\ell = \mathbf{Q}_\ell S_\ell$ with $n \times \ell$ orthonormal $\mathbf{Q}_\ell$ and $\ell \times \ell$ upper triangular S_ℓ, rather than a partial eigenvector decomposition, $\mathbf{A}\mathbf{X}_\ell = \mathbf{X}_\ell \Lambda$ with $n \times \ell$ matrix $\mathbf{X}_\ell$ and $\ell \times \ell$ diagonal Λ. The Schur vectors $\mathbf{q}_1, \ldots, \mathbf{q}_\ell$, i.e., the columns of $\mathbf{Q}_\ell$, span the space of eigenvectors with eigenvalues λ_j if the diagonal elements λ_j of Λ are the diagonal entries of S_ℓ. Detected Schur vectors give an *explicit* way for deflating the detected eigenvectors, or, more precise, for deflating the detected Schur vectors, from the matrix: continue the search with the *deflated matrix* (cf., Exercise 4.3)

$$\widetilde{\mathbf{A}} \equiv (\mathbf{I} - \mathbf{Q}_\ell \mathbf{Q}_\ell^*)\mathbf{A}(\mathbf{I} - \mathbf{Q}_\ell \mathbf{Q}_\ell^*). \quad [4]$$

An eigenvector of $\widetilde{\mathbf{A}}$ is a Schur vector of $\mathbf{A}$.

For many methods, one of the projections is redundant, or more precise, is automatically taken care of. For instance, if $\widetilde{\mathbf{A}}\mathbf{v}$ is the expansion vector and $\mathbf{v} \perp \mathbf{Q}_\ell$, then $(\mathbf{I} - \mathbf{Q}_\ell \mathbf{Q}_\ell^*)\mathbf{v} = \mathbf{v}$ and the projection at the right is not needed.

In Arnoldi's method the projections are not (explicitly) needed when the Schur vectors are kept in the Arnoldi basis as the first ℓ vectors. Because, if $\mathbf{v}_k$ is the last vector in this basis, then $\mathbf{v}_k$ is orthogonal to all preceding basis vectors, in particular, it is already orthogonal to $\mathbf{q}_1, \ldots, \mathbf{q}_\ell$. In Arnoldi's method, the expansion vector $\mathbf{A}\mathbf{v}_k$ is orthogonalised against all preceding basis vectors, in particular, it is orthogonalised against $\mathbf{q}_1, \ldots, \mathbf{q}_\ell$.[5] The Hessenberg matrix in this case has ℓ zero subdiagonal entries: $h_{j+1,j} = 0$ for $j = 1, \ldots, \ell$; the Hessenberg matrix is reduced. Note that keeping the Schur vectors in the Arnoldi basis has the additional advantage that the Schur matrix S_ℓ becomes automatically available (as the left top $\ell \times \ell$ block of $\underline{H}_k$).

Theoretically, deflation is automatically taken care of in Arnoldi's method. The following lemma proves that a Krylov subspace spanned by Arnoldi vectors contains an eigenvector if and only if the corresponding Hessenberg matrix has a subdiagonal element equal to zero.

Lemma 13.10 *Let $\underline{H}_k$ be a $(k+1) \times k$ upper Hessenberg matrix. Assume that*

$$H_k\, y = \vartheta\, y \quad \textit{with } y \textit{ normalized and } \quad h_{k+1,k}\, e_k^* y = 0.$$

Then $h_{j+1,j} = 0$ for a $j < k$. In particular, if $\mathbf{A}\mathbf{V}_k = \mathbf{V}_{k+1}\underline{H}_k$ is an Arnoldi relation, then $\mathbf{A}\mathbf{V}_j = \mathbf{V}_j H_j$, where $\mathbf{V}_j$ is the left $n \times j$ block of $\mathbf{V}_k$ and H_j is the $j \times j$ left upper block of $\underline{H}_k$: $\mathbf{V}_j$ spans an invariant subspace of $\mathbf{A}$ (and contains an eigenvector). If the span of $\mathbf{V}_k$ contains multiples of one eigenvector only, then $h_{2,1} = 0$.

Exercise 13.16. Prove Lemma 13.10.

This suggests a strategy for deflation in Arnoldi's method: if a subdiagonal element of the Hessenberg matrix $\underline{H}_k$ is small, say $|h_{j+1,j}| \leq \varepsilon$, then replace this quantity by 0 and continue the

[4] For stability reasons, the projections have to be performed with a stable version of Gram–Schmidt, as repeated Gram–Schmidt.

[5] Including $\mathbf{q}_1, \ldots, \mathbf{q}_\ell$ in the basis and applying a stable version of Gram–Schmidt for orthogonalising the expansion vector against the basis is more stable than applying a projection $\mathbf{I} - \mathbf{Q}\mathbf{Q}^*$ followed by orthogonalisation against the basis from which $\mathbf{q}_1, \ldots, \mathbf{q}_\ell$ is excluded (which is a kind of 'block' modified Gram–Schmidt).

computational process. Keep $\mathbf{V}_j$ in the process for orthogonalisation, but compute Ritz values only for the $(k+1-j)\times(k-j)$ right lower block of $\underline{H}_k$ and apply the shifted QR steps also only to this right lower block (apply the apropriate rotations also to the $j\times(k-j)$ right upper block of $\underline{H}_k$). The detected eigenvalues and eigenvectors (actually, the space of detected eigenvectors) are **locked**: they are kept in the orthogonalisition steps, to avoid the introduction of spurious eigenvalues, but they are excluded from updating process for approximate eigennpairs.

The QR-algorithm is backward stable (the computed quantities are the exact ones for a slightly perturbed matrix), but it can be forward unstable, specifically if a shift is used that is very close to an eigenvalue: the computed quatities may differ significantly from the ones that would be obtained in exact arithmetic. As a consequence, the size of subdiagonal elements of the Hessenberg matrix may not be reliable. A stable and reliable procedure is obtained by a similarity transformation with a Householder reflection to put the detected eigenvalue at the top left position of $\underline{H}_k$ and then using the `Hessenberg` routine as indicated in Exercise 13.9(e) restore the Hessenberg structure for the $k\times(k-1)$ right bottom other part of the matrix; see the next exercise for details.

Exercise 13.17. **Locking.** Let $\underline{H}_k$ be a $(k+1)\times k$ Hessenberg matrix.
Assume $H_k\,y=\vartheta\,y_k$ such that $|\rho|\le\varepsilon$, where $\rho\equiv h_{k+1,k}\,e_k^*y$.

(a) Show there is a Householder reflection with a normalized k-vector w such that

$$(I_k-2ww^*)e_1=\tau y \quad \text{for some scalar } \tau,\ |\tau|=1,\ \tau e_1^*y\in\mathbb{R}.$$

(b) Show that $(I_{k+1}-2ww^*)\underline{H}_k\,(I_k-2ww^*)$ is of the form

$$(I_{k+1}-2ww^*)\underline{H}_k\,(I_k-2ww^*)=\begin{bmatrix}\vartheta & b^*\\ \rho e_k & \underline{G}_{k-1}\end{bmatrix}$$

Here the w at the left equals the w at the right extended by one 0. The first column of the transformed matrix consists of all zeros except at the first position, where it has the value ϑ, and the last position, where it has the small value ρ. The matrix $\underline{G}_{k-1}$ is (full) $k\times(k-1)$.

(c) Use Exercise 13.9(e) to construct a $k\times k$ unitary matrix Q (as a product of Householder reflections) and a $k\times(k-1)$ upper Hessenberg matrix $\underline{\widetilde{H}}_{k-1}$ such that, with Q° as defined in Exercise 13.9(e),

$$\underline{H}_k\,Q_k=Q_k^\circ\begin{bmatrix}\vartheta & \widetilde{b}^*\\ \rho e_k & \underline{\widetilde{H}}_{k-1}\end{bmatrix} \tag{13.12}$$

(d) Describe a locking procedure based on the above suggestions.

Ideally the filter procedure prevents eigenpairs that are not wanted from being detected. Nevertheless, due to rounding errors (and forward instabilities of the QR-procedure) unwanted eigenpairs are occasionally detected. They can be locked as described above. If memory is an issue and too many unwanted eigenpairs are detected, then it may be desirable to remove them from the Arnoldi relation. Removing locked eigenpairs from the Arnoldi relation is called **purging**. It requires the decoupling of the 'active' Arnoldi vectors from the unwanted locked eigenvector, i.e., $\widetilde{b}^*$ from (13.12) has to be transformed to 0^*. The next lemma tells us that this is possible. Note that if the unwanted eigenvector is to be removed, then there is no need to maintain orthogonality between the unwanted eigenvector and the Arnoldi vectors. The purging procedure is in low dimensional space. It 'sets' $\widetilde{b}^*$ to 0^* using a basis consisting of e_1 and vectors orthogonal to the left eigenvector of H_k with eigenvalue ϑ. The easiest choice for this basis leads to a Hessenberg structure of the remaning part (see Exercise 13.18).

Lemma 13.11 *Let $\underline{H}_k$ be a $(k+1)\times k$ matrix, b a k-vector and ϑ a scalar. There is an orthonormal $(k+1)\times k$ matrix Q and a $(k+1)$-vector y such that*

$$\begin{bmatrix}\vartheta & b^*\\ 0 & \underline{H}_k\end{bmatrix}[e_1,Q]=[e_1,Q^\circ]\begin{bmatrix}\vartheta & 0^*\\ 0 & \underline{\widetilde{H}}_k\end{bmatrix}\quad \textit{and}\quad y^*\begin{bmatrix}\vartheta & b^*\\ 0 & H_k\end{bmatrix}=\vartheta\,y^*. \tag{13.13}$$

The columns of Q form a basis of $y^\perp$, y is a left eigenvector with eigenvalue ϑ.
Here, Q° is the matrix Q expanded with a row of zeros and the column e_{k+2}.

Exercise 13.18. **Purging.** Consider the Arnoldi relation, with locked eigenvector $\mathbf{u}_1$,

$$\mathbf{A}[\mathbf{u}_1, \mathbf{V}_k] = [\mathbf{u}_1, \mathbf{V}_k, \mathbf{v}_{k+1}] \begin{bmatrix} \vartheta & b^* \\ 0 & \underline{H}_k \end{bmatrix}$$

(a) Let $(1, y^{\mathrm{T}})^{\mathrm{T}}$ be a $(k+1)$ left eigenvector as in (13.13). Show that

$$\begin{bmatrix} 1 & -(y^* 0) \\ 0 & I_{k+1} \end{bmatrix} \begin{bmatrix} \vartheta & 0^* \\ 0 & \underline{H}_k \end{bmatrix} = \begin{bmatrix} \vartheta & b^* \\ 0 & \underline{H}_k \end{bmatrix} \begin{bmatrix} 1 & -y^* \\ 0 & I_k \end{bmatrix}$$

(b) Consider the QR-decomposition

$$\begin{bmatrix} -y^* \\ I_k \end{bmatrix} = QR$$

with $(k+1) \times k$ orthonormal Q and $k \times k$ upper triangular R. Show that the columns of the matrix at the left are orthogonal to the left eigenvector $(1, y^{\mathrm{T}})^{\mathrm{T}}$ and conclude that the clumns of Q form an orthonormal basis of the space orthogonal to this left eigenvector.
(c) Put $\widetilde{\underline{H}}_k \equiv R^\circ \underline{H}_k R^{-1}$. Show that $\widetilde{\underline{H}}_k$ is upper Hessenberg and

$$\begin{bmatrix} e_1 & Q^\circ \end{bmatrix} \begin{bmatrix} \vartheta & 0^* \\ 0 & \widetilde{\underline{H}}_k \end{bmatrix} = \begin{bmatrix} \vartheta & b^* \\ 0 & \widetilde{\underline{H}}_k \end{bmatrix} \begin{bmatrix} e_1 & Q \end{bmatrix}$$

(d) Describe a purging procedure based on the above suggestions.

Note that the purged eigenvectors may re-enter the computational process again in a later stage. To prevent this from happening they should be kept in the collection of locked eigenvectors (demanding memory and, since they will be involved then in the orthogonalisation, they contribute to the computational costs).

E Subspace methods

Methods as Arnoldi and Lanczos exploit Krylov subspaces for finding approximate eigenpairs. They can be viewed as subspace variants of the power method. We already learnt that it can be more effective to use shift and invert variants of the power method, multiply by $(\mathbf{A} - \mu\,\mathbf{I})^{-1}$, rather than the (shifted) power method, multiply by $\mathbf{A} - \mu\,\mathbf{I}$, since it leads to amplification of eigenvector components with eigenvalue close to μ. If the shift is fixed in each step, then shift and invert is a power method as well and can be extended to an Arnoldi variant. However, it can be efficient to allow different shifts in subsequent steps of shift and invert. The subspace variant of such an approach does not built Krylov subspaces. This may also be the case if the systems for inverting are not solved exactly. We discuss two non-Krylov subspace methods below, Rational Krylov Sequence (RKS) and Jacobi–Davidson. Of course the advantage of a Krylov structure (with Hessenberg matrices) is lost and the 'orthogonalisation' part in the steps may be more expensive, but less steps may be needed. Note, however, that, if we give up the Krylov structure, restart (and deflation) may be less cumbersome.

Though RKS, the first method that we discuss below, does not built a Krylov subspace, the space is related to a Krylov subspace and leads to structured relations allowing computational savings.

Rational Krylov Sequence method

For a sequence $(\mu_1, \ldots, \mu_n)$ of scalars (in $\mathbb{C}$), consider the polynomials

$$p_k(\zeta) \equiv (\zeta - \mu_1) \cdot \ldots \cdot (\zeta - \mu_k) \quad (\zeta \in \mathbb{C}, k = 1, 2, \ldots, n).$$

The first $k+1$ terms of the **rational Krylov sequence**

$$\mathbf{v}_1, p_1(\mathbf{A})^{-1}\mathbf{v}_1, p_2(\mathbf{A})^{-1}\mathbf{v}_1, p_3(\mathbf{A})^{-1}\mathbf{v}_1, \ldots, p_k(\mathbf{A})^{-1}\mathbf{v}_1$$

span a space of rational functions in $\mathbf{A}$ times $\mathbf{v}_1$:

$$\left\{ \psi(\mathbf{A})\mathbf{v}_1 \,\middle|\, \psi = \frac{q}{p_k} \quad (q \in \mathcal{P}_k) \right\} = \mathcal{K}_{k+1}(\mathbf{A}, p_k(\mathbf{A})^{-1}\mathbf{v}_1).$$

Note that this rational Krylov subspace depends on the **poles** $\mu_1, \ldots, \mu_k$ (whereas a Krylov subspace does not depend on the zeros of the polynomials:

$$\operatorname{span}(\mathbf{v}_1, p_1(\mathbf{A})\mathbf{v}_1, \ldots, p_k(\mathbf{A})\mathbf{v}_1) = \mathcal{K}_{k+1}(\mathbf{A}, \mathbf{v}_1).$$

The rational Krylov subspace incorporates Shift and Invert and requires matrix inversion (i.e., solving linear systems). Since matrix inversion can be expensive, the poles μ_j are often kept the same for a number of consecutive steps. Then the same LU-decomposition of $\mathbf{A} - \mu_j\mathbf{I}$ can be used in these steps.

Consider the **generalised eigenvalue problem**

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{B}\mathbf{x}, \tag{13.14}$$

with eigenpair $(\lambda, \mathbf{x})$, $\mathbf{x} \neq \mathbf{0}$. We assume the **matrix pencil** $\mathbf{A} - \lambda\mathbf{B}$ to be **regular** or non-degenerated, that is, $\det(\mathbf{A} - \mu\mathbf{B})$ is non-zero for some $\mu \in \mathbb{C}$. If a pencil is not regular, then it is said to be **singular**. We allow the eigenvalue λ to have value ∞, which is the case if $\mathbf{x}$ is a kernel vector of $\mathbf{B}$.[6]

The generalised eigenvalue problem can be turned into a standard one if $\mathbf{B}$ or a shifted matrix $\mathbf{A} - \mu\mathbf{B}$ is non-singular: $(\lambda, \mathbf{x})$ is an eigenpair of (13.14) if and only if $(1/(\lambda - \mu), \mathbf{x})$ is an eigenpair of the matrix $(\mathbf{A} - \mu\mathbf{B})^{-1}\mathbf{B}$. Since matrix inversion (i.e., solving systems) is intrinsic for RKS anyway, we discuss this method for the more general problem of generalised eigenvalues rather than standard eigenvalues.

Property 13.12 *The matrix pencil $\mathbf{A} - \lambda\mathbf{B}$ is regular if and only if there are only finitly many $\lambda_j \in \mathbb{C}$, generalised eigenvalues, such that $\mathbf{A} - \lambda_j\mathbf{B}$ is singular.*

Exercise 13.19.

(a) Prove Prop. 13.12.

(b) If $\mathcal{N}(\mathbf{A}) \cap \mathcal{N}(\mathbf{B}) \neq \{\mathbf{0}\}$ then the pencil $\mathbf{A} - \lambda\mathbf{B}$ is singular.

(c) Consider the matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Show that the pencil is singular even though $\mathcal{N}(\mathbf{A}) \cap \mathcal{N}(\mathbf{B}) = \{\mathbf{0}\}$.

[6] To avoid the use of the value ∞, the generalised eigenvalue problem is often formulated as finding a pair (c, s) with $c \in [0, 1]$ and $s \in \mathbb{C}$ such that $c^2 + |s|^2 = 1$ and $c\mathbf{A}\mathbf{x} = s\mathbf{B}\mathbf{x}$. Then $\lambda = s/c$ and, with $c = 0$, $s = 1$, we have the situation where $\lambda = \infty$.

1) Select a $\mu_k \in \mathbb{C}$ and a $\vec{t}_k \in \mathbb{C}^k$ with $e_k^* \vec{t}_k \neq 0$
2) Solve $(\mathbf{A} - \mu_k \mathbf{B})\mathbf{z} = \mathbf{B}(\mathbf{V}_k \vec{t}_k)$ for $\mathbf{z}$
3) $[\mathbf{w}, \vec{h}_k] = \mathtt{Orth}(\mathbf{V}_k, \mathbf{z})$
4) $\eta = \|\mathbf{w}\|_2$, $\mathbf{v}_{k+1} = \mathbf{w}/\eta$, $\mathbf{V}_{k+1} = [\mathbf{V}_k, \mathbf{v}_{k+1}]$

ALGORITHM 13.3. One step RKS (see the text below).

The so-called **QZ-algorithm** (qz in MATLAB) is the method of choice for solving the generalised eigenvalue problems for dense systems. It generalises the QR-algorithm. It computes $n \times n$ unitary matrices $\mathbf{Q}$ and $\mathbf{Z}$ and $n \times n$ upper triangular matrices $\mathbf{S}$ and $\mathbf{T}$ such that

$$\mathbf{AQ} = \mathbf{ZS} \quad \text{and} \quad \mathbf{BQ} = \mathbf{ZT}.$$

The eigenvalues λ are on the diagonal of $\mathbf{T}^{-1}\mathbf{S}$.

If the pencil is singular, then the singular part induces an "unstable eigenvalue" $(\varepsilon_A/\varepsilon_B)$, while the QZ-algorithm is stable, representing the singular part as small common diagonal entries of $\mathbf{S}$ and $\mathbf{T}$, say ε_A and ε_b, respectively.

For ease of notation, we assume that $\mathbf{A}$ is non-singular: if $\mathbf{A}$ is singular, consider the shifted problem $(\mathbf{A} - \mu\mathbf{B})\mathbf{x} = (\lambda - \mu)\mathbf{B}\mathbf{x}$ with $\mathbf{A} - \mu\mathbf{B}$ non-singular.

For computing eigenpairs $(\lambda, \mathbf{x})$ of (13.14), the **rational Krylov sequence** method (**RKS**) computes an orthonormal matrix $\mathbf{V}_{k+1}$ that spans the space spanned by $\mathbf{w}_0, \ldots, \mathbf{w}_k$ with

$$\mathbf{w}_0 \equiv \mathbf{v}_1, \quad \mathbf{w}_j \equiv (\mathbf{A} - \mu_j \mathbf{B})^{-1}\mathbf{B}\mathbf{w}_{j-1} \qquad (j = 1, \ldots, k).$$

or, equivalently,

$$\operatorname{span}(\mathbf{V}_{k+1}) = \mathcal{K}_{k+1}(\mathbf{A}^{-1}\mathbf{B}, \widetilde{p}_k(\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{v}_1), \tag{13.15}$$

where $\widetilde{p}(\zeta) \equiv (1 - \mu_1\zeta)\ldots(1 - \mu_k\zeta) \quad (\zeta \in \mathbb{C})$.

Exercise 13.20.

(a) Prove (13.15).

(b) If $\mathbf{B}$ is non-singular, then

$$\mathcal{K}_{k+1}(\mathbf{A}^{-1}\mathbf{B}, \widetilde{p}_k(\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{v}_1) = \mathcal{K}_{k+1}(\mathbf{B}^{-1}\mathbf{A}, p_k(\mathbf{B}^{-1}\mathbf{A})^{-1}\mathbf{v}_1)$$

(c) If $\mu_j = \mu$ for all $j = 1, 2, \ldots$, then

$$\mathcal{K}_{k+1}(\mathbf{A}^{-1}\mathbf{B}, \widetilde{p}_k(\mathbf{A}^{-1}\mathbf{B})^{-1}\mathbf{v}_1) = \mathcal{K}_{k+1}((\mathbf{A} - \mu\mathbf{B})^{-1}\mathbf{B}, \mathbf{v}_1).$$

Prove that, in this case and choosing $t_k = e_k$, RKS is equivalent to Arnoldi for the matrix $(\mathbf{A} - \mu\mathbf{B})^{-1}\mathbf{B}$.

Assume $\mathbf{V}_k$ is available and $\mathbf{v}_{k+1}$ is computed as in ALG. 13.3. If $\underline{H}_{k-1}$ and $\underline{K}_{k-1}$ are $k \times (k-1)$ upper Hessenberg matrices such that

$$\mathbf{A}\mathbf{V}_k \underline{H}_{k-1} = \mathbf{B}\mathbf{V}_k \underline{K}_{k-1},$$

then this **RKS relation** is expanded to $\mathbf{A}\mathbf{V}_{k+1} \underline{H}_k = \mathbf{B}\mathbf{V}_{k+1} \underline{K}_k$ by

$$\mathbf{V}_{k+1} = [\mathbf{V}_k, \mathbf{v}_{k+1}], \quad \underline{H}_k = \begin{bmatrix} \underline{H}_{k-1} & \vec{h}_k \\ 0 & \eta \end{bmatrix}, \qquad \underline{K}_k = \begin{bmatrix} \underline{K}_{k-1} & \vec{h}_k \mu_k + \vec{t}_k \\ 0 & \eta \mu_k \end{bmatrix}.$$

$(\mathbf{V}_{k+1}, \underline{H}_k, \underline{K}_k)$ is an **RKS tripel** of order $k+1$ for the pencil $\mathbf{A} - \lambda\mathbf{B}$.

Approximate eigenvalues ϑ_j with approximate eigenvectors $\mathbf{u}_j$ for the generalised eigenvalue problem are obtained as

$$\mathbf{u}_j = \mathbf{V}_k\, H_k\, y_j \quad \text{where } (\vartheta_j, y_j) \text{ solves} \quad \vartheta_j H_k\, y_j = K_k\, y_j.$$

The k-dimensional generalized eigenvalue problem can be solved via the QZ-algorithm. The approximate eigenpairs $(\vartheta_j, \mathbf{u}_j)$ can be viewed as being obtained by projecting (13.14) onto $\mathcal{V}_k \equiv \text{span}(\mathbf{V}_K)$, as we will see below in Exercise 13.21(b). With residual $\mathbf{r}_j \equiv \mathbf{A}\mathbf{u}_j - \vartheta_j \mathbf{B}\mathbf{u}_j$ we have that

$$\mathbf{r}_j = -\eta(\mathbf{A} - \mu_k \mathbf{B})\mathbf{v}_{k+1} e_k^* y_j. \tag{13.16}$$

Then progress can be monitored by $|\eta\, e_k^* y_j|$.

Exercise 13.21.

(a) Prove (13.16).

(b) Suppose H_k is non-singular. Show that in this approach, $(1/(\vartheta_j - \mu_k), \mathbf{u}_j)$ is an Ritz pair for $(\mathbf{A} - \mu_k \mathbf{B})^{-1}\mathbf{B}$ obtained from the search subspace $\mathcal{V}_k$ (and tested against $\mathcal{V}_k$ as well).

RKS shares advantages of Arnoldi: the orthogonalisation scalars are used in the formation of the projected problem. The fact that the projected matrices are upper Hessenberg is exploited. RKS is the rational Krylov variant of Arnoldi. Implicit restart procedure that preserve the Hessenberg structure at restart have been designed.

Jacobi–Davidson method

Jacobi–Davidson (JD) also obtains expansion vectors from a shift and invert approach (as RKS in step 2 of Alg. 13.3). However, rather than formulating the shift and invert step such that the subsequent orthogonalisation and projection steps are as efficient as possible as in RKS, JD aims for optimal expansion using a well-conditioned proper variant of the shift and invert system, called the JD correction equation. This leads to faster convergence and allows inexact solves of the 'shift and invert system', i.e., preconditioned iterative linear systems solvers can be used to solve the correction equation to some degree. Often a modest degree of accuracy already leads to fast convergence.

For high dimensional problems (n is huge), exact shift and invert is not feasible and this is where the main advantage of the JD approach comes in. JD allows to use preconditioners and (inaccurate) solutions obtained with (a few steps of) some iterative linear solver.

Exercise 13.22. **Jacobi-Davidson.** Consider (13.14) with regular pencil.

Suppose we have an approximate eigenvector $\mathbf{u}$. Select

$$\vartheta \equiv \frac{\mathbf{u}^*\mathbf{A}\mathbf{u}}{\mathbf{u}^*\mathbf{B}\mathbf{u}}$$

as approximate eigenvalue. We are interested in finding corrections $\varepsilon \in \mathbb{C}$ to ϑ and $\mathbf{t} \in \mathbb{C}^n$, $\mathbf{t} \perp \mathbf{u}$, to $\mathbf{u}$ such that $\lambda = \vartheta + \varepsilon$ is the wanted eigenvalue with eigenvector $\mathbf{x} = \mathbf{u} + \mathbf{t}$. Put $\mathbf{v} \equiv \mathbf{B}\mathbf{u}$. Let $\mathbf{r} \equiv \mathbf{A}\mathbf{u} - \vartheta\mathbf{B}\mathbf{u}$ be the residual.

(a) Show that $\mathbf{r} \perp \mathbf{u}$.

(b) Show that $\mathbf{0} = \mathbf{r} + \varepsilon\mathbf{v} + (\mathbf{A} - \vartheta\mathbf{B})\mathbf{t} - \varepsilon\mathbf{B}\mathbf{t}$.

(c) Neglecting $\mathcal{O}(|\varepsilon|\,\|\mathbf{t}\|_2)$-terms ('small'^{2} terms), implies that

$$(\mathbf{I} - \mathbf{v}(\mathbf{u}^*\mathbf{v})^{-1}\mathbf{u}^*)(\mathbf{A} - \vartheta\mathbf{B})\mathbf{t} = -\mathbf{r}, \quad \text{where} \quad \mathbf{t} \perp \mathbf{u}. \tag{13.17}$$

This is the **JD correction equation**. Note that, in case of the standard eigenvalue problem, i.e., $\mathbf{B} = \mathbf{I}$, and $\mathbf{u}$ is normalised, this correction equation reduces to

$$(\mathbf{I} - \mathbf{u}\mathbf{u}^*)(\mathbf{A} - \vartheta\,\mathbf{I})\mathbf{t} = -\mathbf{r}, \quad \text{where} \quad \mathbf{t} \perp \mathbf{u}. \tag{13.18}$$

(d) Put $\widetilde{\mathbf{A}} \equiv (\mathbf{I} - \mathbf{v}(\mathbf{u}^*\mathbf{v})^{-1}\mathbf{u}^*)(\mathbf{A} - \vartheta\mathbf{B})$. Prove that $\mathcal{K}_s(\widetilde{\mathbf{A}}, \mathbf{r}) \perp \mathbf{u}$.

JACOBI–DAVIDSON

%% *Initiation*

Select an $n \times \ell$ orthonormal matrix $\mathbf{V}$.

Compute $\mathbf{W} \equiv \mathbf{AV}$ and the interaction matrix $H = \mathbf{V}^*\mathbf{W}$.

repeat

  %% *Extraction*

  1) Compute the Schur form $H = USU^*$.

     Order the Schur form $H = USU^*$.

  2) Select $(\vartheta, y) = (S_{1,1}, Ue_1)$.

  3) Compute $\mathbf{u} = \mathbf{V}_k y$ and the residual $\mathbf{r} = \mathbf{W}y - \vartheta\mathbf{u}$.

  %% *Stopping criterion*

  4) If $\|\mathbf{r}\|_2 \leq tol$, $\mathbf{x} = \mathbf{u}$, $\lambda = \vartheta$, break, end

  %% *Expansion*

  5) Solve (approximately) for $\mathbf{t} \perp \mathbf{u}$

$$(\mathbf{I} - \mathbf{uu}^*)(\mathbf{A} - \vartheta\mathbf{I})(\mathbf{I} - \mathbf{uu}^*)\mathbf{t} = -\mathbf{r}.$$

  6) $\mathbf{v} = \mathtt{Orth}(\mathbf{V}, \mathbf{t})$.

  7) $\mathbf{V} \leftarrow [\mathbf{V}, \mathbf{v}]$, $H \leftarrow [H; \mathbf{v}^*\mathbf{W}]$

  8) $\mathbf{w} = \mathbf{Av}$, $\mathbf{W} \leftarrow [\mathbf{W}, \mathbf{w}]$, $H \leftarrow [H, \mathbf{V}^*w]$, $\ell \leftarrow \ell + 1$.

end repeat

ALGORITHM 13.4. This is the basic Jacobi–Davidson scheme for computing one approximate eigenpair $(\lambda, \mathbf{x})$ with residual accuracy *tol* of the standard eigenvalue problem $\mathbf{Ax} = \lambda\mathbf{x}$. The scheme here uses Ritz pairs and an orthonormal basis of the search subspace $\mathcal{V}$ spanned by $\mathbf{V}$. The expansion vector $\mathbf{t}$ of $\mathcal{V}$ is obtained as an (approximate) solution of the JD correction equation and is orthonormalised against $\mathbf{V}$. Substep 5 is essential for Jacobi–Davidson. All subspace methods basically follow the above scheme except for the way they compute an expansion vector $\mathbf{t}$ of the search subspace.

This shows that Krylov methods as GMRES and Bi-CGSTAB are suited for obtaining approximate solutions $\mathbf{t}$ of the JD correction equation. If the initial guess $\mathbf{t}_0$ is $\mathbf{0}$, then all iterates are automatically orthogonal to $\mathbf{u}$. Often an effective expansion vector $\mathbf{t}$ is obtained with s steps of preconditioned GMRES.

The correction equations are solved to some accuracy and the solutions are used to expand the search subspace, see ALG. 13.4. If the correction equation is solved exactly, JD asymptotically converges quadratically. When solved inaccurately, convergence is usually still fast.

Jacobi-Davidson tends to converge quickly to some eigenpair. In order to avoid convergence to an unwanted eigenpair it is helpfull to have a "thick" start. As a start a number of Arnoldi steps is recommended.

The scheme in ALG. 13.4 can be extended with a (thick) restart by extending substep 1:

1.c) if $\ell = k_{\max}$, then %% *restart*

$$\ell \leftarrow k_{\min},$$

$$\mathbf{V} \leftarrow \mathbf{V}U(:, 1:\ell), \quad \mathbf{W} \leftarrow \mathbf{W}U(:, 1:\ell), \quad H \leftarrow S(1:\ell, 1:\ell).$$

And, if more than one eigenvector (Schur vector) is required, then it can be extended with deflation by modifying step 4:

4) if $\|\mathbf{r}\|_2 < tol$, then %% *lock a Schur vector*
$\mathbf{Q} \leftarrow [\mathbf{Q}, \mathbf{u}]$,
if the number of detected Schur vectors (the columns of $\mathbf{Q}$) is as desired
stop
else %% *deflate*
$\mathbf{V} \leftarrow \mathbf{V}U(:, 2:\ell)$, $\mathbf{W} \leftarrow \mathbf{W}U(:, 2:\ell)$, $H \leftarrow S(2:\ell, 2:\ell)$,
$\ell \leftarrow \ell - 1$, and go to substep 3.

If deflation is included, the correction equation in substep 5 should be deflated as well:

$$(\mathbf{I} - [\mathbf{Q}, \mathbf{u}][\mathbf{Q}, \mathbf{u}]^*)(\mathbf{A} - \vartheta\mathbf{I})(\mathbf{I} - [\mathbf{Q}, \mathbf{u}][\mathbf{Q}, \mathbf{u}]^*)\mathbf{t} = -\mathbf{r}, \qquad \mathbf{t} \perp [\mathbf{Q}, \mathbf{u}].$$

Exercise 13.23. Consider the Jacobi–Davidson scheme ALG. 13.4.

(a) Adapt the scheme to harmonic Ritz values.

(b) Adapt the scheme for solving the generalised eigenvalue problem. Include the adaptation also in the restart and deflation.

To compare JD with **Davidson**,

$$\text{solve } \mathbf{M}\mathbf{t} = -\mathbf{r} \text{ for } \mathbf{t},$$

and inexact Rayleigh Quotient Iteration with subspace acceleration (aiRQI),

$$\text{solve } (\mathbf{A} - \vartheta\mathbf{B})\mathbf{t} = \mathbf{u},$$

suppose $\mathbf{M}$ is an $n \times n$ matrix that approximates $\mathbf{A} - \vartheta\mathbf{B}$ in some sense and such that systems of the form $\mathbf{M}\mathbf{y} = \mathbf{z}$ are ease to solve for $\mathbf{y}$. $\mathbf{M}$ can be preconditioner as an ILU-decomposition of $\mathbf{A} - \tau\mathbf{B}$ or $\mathbf{M}^{-1}$ can represent $q(\mathbf{A} - \vartheta\mathbf{B})\mathbf{z}$, with q a polynomial of degree $< s$ as obtained by applying a s steps of a Krylov method to solve $(\mathbf{A} - \vartheta\mathbf{B})\mathbf{y} = \mathbf{z}$ (then the residual is equal to $(\mathbf{I} - (\mathbf{A} - \vartheta\mathbf{B})q(\mathbf{A} - \vartheta\mathbf{B}))\mathbf{z}$), or a combination of both, $\mathbf{K}^{-1}q((\mathbf{A} - \vartheta\mathbf{B})\mathbf{K}^{-1})\mathbf{z}$, a Krylov solver, preconditioned with $\mathbf{K}$.

Now consider the correction equation that arises by replacing $\mathbf{A} - \vartheta\mathbf{B}$ by $\mathbf{M}$ in the JD correction equation (13.17):

$$(\mathbf{I} - \mathbf{v}(\mathbf{u}^*\mathbf{v})^{-1}\mathbf{u}^*)\mathbf{M}\mathbf{t} = -\mathbf{r}, \qquad \mathbf{t} \perp \mathbf{u}. \tag{13.19}$$

Property 13.13 *If* $\mathbf{w}$ *solves* $\mathbf{M}\mathbf{w} = \mathbf{v}$ *and* $\mathbf{s}$ *solves* $\mathbf{M}\mathbf{s} = -\mathbf{r}$*, then* (13.19) *is solved by*

$$\mathbf{t} \equiv \left(\mathbf{I} - \mathbf{w}(\mathbf{u}^*\mathbf{w})^{-1}\mathbf{u}^*\right)\mathbf{s}.$$

This result is of interest in itself: it gives a way to incorporate a preconditioner in a Krylov solver for the JD correction equation , see Exercise 13.24. But it also helps to understand the advantage of JD over Davidson and over aiRQI, see Exercise 13.25.

Exercise 13.24. **Preconditioning the JD correction equation.**

(a) Prove Property 13.13.

Apparently, if $\mathbf{M}$ is a preconditioner for $\mathbf{A} - \vartheta\mathbf{B}$, then $\left(\mathbf{I} - \mathbf{w}(\mathbf{u}^*\mathbf{w})^{-1}\mathbf{u}^*\right)\mathbf{M}^{-1}$ can be viewed as a preconditioner for $(\mathbf{I} - \mathbf{v}(\mathbf{u}^*\mathbf{v})^{-1}\mathbf{u}^*)(\mathbf{A} - \vartheta\mathbf{B})(\mathbf{I} - \mathbf{u}\mathbf{u}^*)$ (why?).

(b) Prove that multiplication of the projected matrix followed by a preconditioning step equals

$$\left(\mathbf{I} - \mathbf{w}(\mathbf{u}^*\mathbf{w})^{-1}\mathbf{u}^*\right)\mathbf{M}^{-1}(\mathbf{A} - \vartheta\mathbf{B})(\mathbf{I} - \mathbf{u}\mathbf{u}^*)$$

and conclude that, starting with $\mathbf{t}_0 = \mathbf{0}$ it suffices to multiply by

$$\left(\mathbf{I} - \mathbf{w}(\mathbf{u}^*\mathbf{w})^{-1}\mathbf{u}^*\right)\mathbf{M}^{-1}(\mathbf{A} - \vartheta\mathbf{B})$$

in a preconditioned Krylov solver (as GMRES).

Exercise 13.25. **JD versus Davidson and aiRQI.**
For ease of discussion we assume $\mathbf{B} = \mathbf{I}$. Suppose $\mathbf{A} - \vartheta\mathbf{I}$ and $\mathbf{M}$ are of the form

$$\mathbf{A} - \vartheta\mathbf{I} = \begin{bmatrix} 0 & \mathbf{a}^* \\ \mathbf{r} & \mathbf{B} \end{bmatrix} \quad \text{and} \quad \mathbf{M} = \begin{bmatrix} \mu & \mathbf{b}^* \\ \mathbf{s} & \mathbf{B} \end{bmatrix}, \quad \text{respectively.}$$

Note that $\mathbf{A} - \vartheta\mathbf{I}$ takes this form with respect to an orthonormal basis with $\mathbf{u}$ as first vector and ϑ as the Rayleigh quotient for $\mathbf{u}$. Here, we modelled the assumption that $\mathbf{M}$ is an approximation of $\mathbf{A} - \vartheta\mathbf{I}$ with perturbation term Δ ($\mathbf{M} = \mathbf{A} - \vartheta\mathbf{I} + \Delta$) that is small as compared to $\|\mathbf{A} - \vartheta\mathbf{I}\|_2$. If $\mathbf{B}$ is well-conditioned (which will be the case if λ is simple and ϑ is close to λ) then the block of Δ that corresponds to the block $\mathbf{B}$ is negligible, also as compared to $\mathbf{B}^{-1}$. Therefore, to ease the discussion, we neglected this block. Since 0 and $\mathbf{r}$ (and $\mathbf{a}$ if $\mathbf{A}$ is Hermitian) are small, the corresponding blocks of Δ can not be negelected.

(a) Show that the vector

$$\mathbf{B}^{-1}\mathbf{r} \tag{13.20}$$

characterizes the expansion vector obtained with the exact solution of the Jacobi–Davidson correction equation.

(b) Show that

$$\mathbf{M} = \begin{bmatrix} 1 & \mathbf{b}^*\mathbf{B}^{-1} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \chi & \mathbf{0}^* \\ \mathbf{0} & \mathbf{B} \end{bmatrix} \begin{bmatrix} 1 & \mathbf{0}^* \\ \mathbf{B}^{-1}\mathbf{s} & \mathbf{I} \end{bmatrix} \quad \text{with} \quad \chi \equiv \mu - \mathbf{b}^*\mathbf{B}^{-1}\mathbf{s};$$

χ is the Schur complement. Conclude that

$$\mathbf{M}^{-1} \begin{bmatrix} 1 \\ \mathbf{0} \end{bmatrix} = \frac{1}{\chi} \begin{bmatrix} 1 \\ -\mathbf{B}^{-1}\mathbf{s} \end{bmatrix} \quad \text{and} \quad \mathbf{M}^{-1} \begin{bmatrix} 0 \\ \mathbf{r} \end{bmatrix} = \begin{bmatrix} \frac{-\mathbf{b}^*\mathbf{B}^{-1}\mathbf{r}}{\chi} \\ \mathbf{B}^{-1}\left(\mathbf{r} + \frac{\mathbf{b}^*\mathbf{B}^{-1}\mathbf{r}}{\chi}\mathbf{s}\right) \end{bmatrix}.$$

Hence, with an inexact Rayleigh Quotient step, the search subspace is essentially expanded by

$$\mathbf{B}^{-1}\mathbf{s}, \tag{13.21}$$

while

$$\mathbf{B}^{-1}\left(\mathbf{r} + \frac{\mathbf{b}^*\mathbf{B}^{-1}\mathbf{r}}{\chi}\mathbf{s}\right) \tag{13.22}$$

represents the expansion vector for Davidson and (13.20) for Jacobi–Davidson. Why?

(c) Compare the quality of the expansion vectors. First consider the case where both $\mathbf{A}$ and $\mathbf{M}$ are Hermitian (then $\mathbf{a} = \mathbf{r}$ and $\mathbf{b} = \mathbf{s}$) and both $|\mu|$ and $\|\mathbf{s}\|_2$ are of order $\|\mathbf{r}\|_2$.

F Higher order eigenvalue problems.

A **quadratic eigenvalue problem** is of the form

$$\mathbf{A}\mathbf{x} + \lambda\mathbf{C}\mathbf{x} + \lambda^2\mathbf{M}\mathbf{x} = \mathbf{0}. \tag{13.23}$$

Here, $\mathbf{A}, \mathbf{C}, \mathbf{M}$ are given $n \times n$ matrices. We have to find an **eigenvector** $\mathbf{x}$, that is, a non-trivial n-vector, and the associated **eigenvalue** λ, a complex number.

In an important class of applications comes from partial differential problems. Then, $\mathbf{A}$ is the discretization of a partial differential operator. $\mathbf{A}$ is often Hermitian. $\mathbf{C}$ represents the boundary conditions. It is often anti-Hermitian, but its rank is relatively low (if the problem is a discretised 3-d partial differential problem, then the rank of $\mathbf{C}$ is of order $\sqrt[2/3]{n}$: $n = n_x n_y n_z$ and the rank of $\mathbf{C}$ is proportional to $n_x n_y$). $\mathbf{M}$ is the mass matrix, that is, the discretization of the identity operator. $\mathbf{M}$ is positive definite.

If $\mathbf{M}$ is positive definite, then a Cholesky decomposition might be helpful (if it is feasible to construct): if $\mathbf{M} = \mathbf{L}\mathbf{L}^*$, then (13.23) can be simplified to $\widetilde{\mathbf{A}}\widetilde{\mathbf{x}} + \lambda\widetilde{\mathbf{C}}\widetilde{\mathbf{x}} + \lambda^2\widetilde{\mathbf{x}} = \mathbf{0}$, where $\widetilde{\mathbf{A}} \equiv \mathbf{L}^{-1}\mathbf{A}\mathbf{L}^{-*}$ and $\widetilde{\mathbf{C}} \equiv \mathbf{L}^{-1}\mathbf{C}\mathbf{L}^{-*}$. What is the relation between $\widetilde{\mathbf{x}}$ and $\mathbf{x}$.

The quadratic eigenvalue problem can be **linearised** to generalised eigenvalue problem:

$$\begin{bmatrix} \mathbf{C} & \mathbf{A} \\ \mathbf{I} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \lambda\mathbf{x} \\ \mathbf{x} \end{bmatrix} + \lambda \begin{bmatrix} \mathbf{M} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \lambda\mathbf{x} \\ \mathbf{x} \end{bmatrix}. \tag{13.24}$$

Exercise 13.26. **Quadratic eigenvalue problems.**

(a) Prove that any solution of (13.23) gives a solution of (13.24) and the other way around.

(b) The linearisation is not unique. If $\mathbf{A}$ is non-singular, then we can replace the identity block $\mathbf{I}$ in (13.24) by $\mathbf{A}$. But also by $-\mathbf{A}^*$, etc..

Exercise 13.27. **Deflation.** In some sense, we generalise in this exercise the Hotelling deflation of Lecture 4 to polynomial eigenvalue problems. We explain how to shift a detected eigenvalue to a non-dominant position (actually to any position that is suitable).

(a) For $\mathbf{y} \in \mathbb{C}^n$, $\|\mathbf{y}\|_2 = 1$ and $\alpha \in \mathbb{C}, \alpha \neq 0$, consider

$$\mathbf{T}_\alpha \equiv \mathbf{T}[\alpha] \equiv \mathbf{I} - \tfrac{1}{\alpha}\mathbf{y}\,\mathbf{y}^*.$$

Prove that $\mathbf{T}_\alpha$ is non-singular if $\alpha \neq 1$ and $\mathbf{T}_\alpha^{-1} = \mathbf{T}_{1-\alpha}$.

(b) For $n \times n$ matrices $\mathbf{A}_0, \ldots, \mathbf{A}_\ell$, consider the **matrix polynomial**

$$\mathbf{P}(\lambda) \equiv \mathbf{A}_\ell\lambda^\ell + \mathbf{A}_{\ell-1}\lambda^{\ell-1} + \ldots + \mathbf{A}_1\lambda + \mathbf{A}_0 \qquad (\lambda \in \mathbb{C}).$$

(c) Show that generalised eigenvalue problems and quadratic eigenvalue problems are polynomial eigenvalue problems.

Consider a **polynomial eigenpair** $(\lambda_1\mathbf{x}_1)$ of $\mathbf{P}$ with $\mathbf{x}_1$ normalised, that is,

$$\lambda_1 \in \mathbb{C}, \quad \mathbf{x}_1 \in \mathbb{C}^n, \|\mathbf{x}_1\|_2 = 1, \quad \text{and} \quad \mathbf{P}(\lambda_1)\mathbf{x}_1 = \mathbf{0}.$$

Select a $\tau \in \mathbb{C}$, $\tau \neq \lambda_1$. Consider the 'deflated' matrix polynomial

$$\widetilde{\mathbf{P}}(\lambda) \equiv \mathbf{P}(\lambda)\,\left(\mathbf{I} - \tfrac{\tau-\lambda_1}{\lambda-\lambda_1}\,\mathbf{x}_1\,\mathbf{x}_1^*\right) \qquad (\lambda \in \mathbb{C}).$$

(d) Relate $\mathbf{I} - \frac{\tau-\lambda_1}{\lambda-\lambda_1}\mathbf{x}_1\mathbf{x}_1^*$ to a $\mathbf{T}_\alpha$.

(e) Prove that $\widetilde{\mathbf{P}}$ is a matrix polynomial of degree ℓ.

(f) Consider an eigenpair $(\lambda_i, \mathbf{x}_i)$ of $\mathbf{P}$. Define

$$\widetilde{\mathbf{x}}_i \equiv \left(\mathbf{I} - \frac{\tau-\lambda_1}{\tau-\lambda_i}\,\mathbf{x}_1\,\mathbf{x}_1^*\right)\mathbf{x}_i \qquad (i = 2, 3, \ldots).$$

Prove that $(\tau, \mathbf{x}_1)$ and $(\lambda_i, \widetilde{\mathbf{x}}_i)$ $(i = 2, 3, \ldots)$ are eigenpairs of $\widetilde{\mathbf{P}}$.
Explain how $\mathbf{x}_i$ can be obtained from $\widetilde{\mathbf{x}}_i$.

(g) Relate the above deflation strategy to the Hotelling deflation of Lecture 4.

Consider the higher degree (degree ℓ) **polynomial eigenvalue problem**

$$\mathbf{x} - \lambda\,\mathbf{M}_1\,\mathbf{x} - \lambda^2\,\mathbf{M}_2\,\mathbf{x} - \ldots - \lambda^\ell\,\mathbf{M}_\ell\,\mathbf{x} = \mathbf{0},^{7} \tag{13.25}$$

[7] For convenience, the matrix $\mathbf{M}_0$ has been normalized (by multiplying by $\mathbf{M}_0^{-1}$, possibly after shifting λ if $\mathbf{M}_0$ is singular).

where $\mathbf{M}_i$ are give $n \times n$ matrices. If $\lambda \in \mathbb{C}$ and $\mathbf{x} \in \mathbb{C}^n$ satisfies (13.25), then $(\lambda, \mathbf{x})$ is an eigenpair of the polynomial problem (13.25).

The problem can be reformulated into a linear one, using a companion formulation as in (13.24). The following observation gives an aalternative. It formulates the problem as a linear one on the space of smooth n-vector valued functions. The linear operator is an integraloperator.

For smooth vector-valued functions $\mathbf{f} : \mathbb{C} \to \mathbb{C}^n$, define the operator $\mathbb{A}$ by

$$\mathbb{A}\mathbf{f}(\zeta) \equiv \int_0^\zeta \mathbf{f}(\vartheta)\ \mathrm{d}\vartheta + \sum_{i=1}^{\ell} \mathbf{M}_i\, \mathbf{f}^{(i-1)}(0) \qquad (\zeta \in \mathbb{C}).$$

Note that $\mathbb{A}\mathbf{f}$ is in $C^{(\infty)}(\mathbb{C}, \mathbb{C}^n)$ if $\mathbf{f}$ is in this space. Moreover $\mathbb{A}$ is linear, whence for a scalar $\mu \in \mathbb{C}$, $(\mu, \mathbf{f})$ is an eigenpair iff $\mathbb{A}\mathbf{f} = \mu\mathbf{f}$. Differentiation of the eigenvalue equation leads to the differential equation $\mathbf{f}' = \frac{1}{\mu}\mathbf{f}$.

Theorem 13.14 *For $\lambda \in \mathbb{C}$ and $\mathbf{x} \in \mathbb{C}^n$, define*

$$\mathbf{f}(\zeta) \equiv \mathbf{x} \exp(\zeta\lambda) \qquad (\zeta \in \mathbb{C}). \tag{13.26}$$

If $(\lambda, \mathbf{x})$ is an eigenpair of (13.25), then $(\frac{1}{\lambda}, \mathbf{f})$ is an eigenpair of $\mathbb{A}$.

Conversely, if $\lambda \in \mathbb{C}$ and $\mathbf{f}$ is an analytic n-vector valued function such that $(\frac{1}{\lambda}, \mathbf{f})$ is an eigenpair of $\mathbb{A}$, then $\mathbf{f}$ is of the form (13.26) and $(\lambda, \mathbf{x})$ is an eigenpair of (13.25).

Exercise 13.28. ***Proof of Theorem*** 13.14**.** Prove the theorem.

The definition of eigenpair can be generalized to invariant subspace: an $n \times p$ matrix $\mathbf{X}$ is an invariant subspace of (13.25) if, for some $p \times p$ matrix S we have that

$$\mathbf{X} - \mathbf{M}_1\, \mathbf{X}\, S - \mathbf{M}_2\, \mathbf{X}\, S^2 - \ldots - \mathbf{M}_\ell\, \mathbf{X}\, S^\ell = \mathbf{0}. \tag{13.27}$$

Exercise 13.29. Relate eigenpairs of S to eigenpairs of (13.25).

Using the expression $\mathbf{F}(\zeta) = \mathbf{X} \exp(\zeta S)$, the theorem can be formulated for invariant subspaces.

Exploiting these equivalences requires a formulation of the methods dicussed sofar for vector-valued function spaces (what basis?) rather than vectors.

The above aproach can be used te reformulate analytic eigenvalue problems for vectors to linear ones for smooth functions.

Let $\zeta \rightsquigarrow \mathbf{M}(\zeta)$ be an analytic map from $\mathbb{C}$ (or some neighborhood of 0 in $\mathbb{C}$) to the space of $n \times n$ matrices. We are interested in solving the **analytic eigenvalue problem**, that is, in finding eigenpairs $(\lambda, \mathbf{x})$ for which $\mathbf{x} - \lambda\mathbf{M}(\lambda)\mathbf{x} = \mathbf{0}$. The associated integral operator can be defined by

$$\mathbb{A}\mathbf{f}(\zeta) \equiv \int_0^\zeta \mathbf{f}(\vartheta)\ \mathrm{d}\vartheta + \sum_{i=0}^{\infty} \frac{1}{i!}\, \mathbf{M}^{(i)}(0)\, \mathbf{f}^{(i)}(0) \qquad (\zeta \in \mathbb{C})$$

for n-vector valued functions $\mathbf{f}$ that are analytic on some neighborhood of $\mathbf{0}$ in $\mathbb{C}$ and for which the infinite sum makes sense.

Exercise 13.30. Check that this generalises the suggested approach for polynomial eigenvalue problems.