Chapter 10
Eigenvalue Problems

Eigenvalue problems are omnipresent in physics. Important examples are the time independent Schrödinger equation in a finite orthogonal basis (Chap. 10)

$$\sum_{j=1}^{M} < \phi_{j'}|H|\phi_j > C_j = EC_{j'} \tag{10.1}$$

or the harmonic motion of a molecule around its equilibrium structure (Sect. 15.4.1)

$$\omega^2 m_i(\xi_i - \xi_i^{eq}) = \sum_j \frac{\partial^2 U}{\partial \xi_i \partial \xi_j}(\xi_j - \xi_j^{eq}). \tag{10.2}$$

Most important are ordinary eigenvalue problems,[1] *which involve the solution of a homogeneous system of linear equations*

$$\sum_{j=1}^{N} a_{ij}x_j = \lambda x_i \tag{10.3}$$

with a Hermitian (or symmetric, if real) matrix [113]

$$a_{ji} = a_{ij}^*. \tag{10.4}$$

The couple $(\lambda, \boldsymbol{x})$ *consisting of an eigenvector* $\boldsymbol{x}$ *and the corresponding eigenvalue* λ *is called an eigenpair.*

[1] *We do not consider general eigenvalue problems here.*

P.O.J. Scherer, *Computational Physics*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-61088-7_10

Matrices of small dimension can be diagonalized directly by determining the roots of the characteristic polynomial and solving a homogeneous system of linear equations. The Jacobi method uses successive rotations to diagonalize a matrix with a unitary transformation. A very popular method for not too large symmetric matrices reduces the matrix to tridiagonal form which can be diagonalized efficiently with the QL algorithm. Some special tridiagonal matrices can be diagonalized analytically. Special algorithms are available for matrices of very large dimension, for instance the famous Lanczos method.

10.1 Direct Solution

For matrices of very small dimension (2, 3) the determinant

$$\det \left| a_{ij} - \lambda \delta_{ij} \right| = 0 \tag{10.5}$$

can be written explicitly as a polynomial of λ. The roots of this polynomial are the eigenvalues. The eigenvectors are given by the system of equations

$$\sum_j (a_{ij} - \lambda \delta_{ij}) u_j = 0. \tag{10.6}$$

10.2 Jacobi Method

Any symmetric 2×2 matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \tag{10.7}$$

can be diagonalized by a rotation of the coordinate system. Rotation by the angle φ corresponds to an orthogonal transformation with the rotation matrix

$$R_\varphi = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix}. \tag{10.8}$$

In the following we use the abbreviations

$$c = \cos\varphi, \quad s = \sin\varphi, \quad t = \tan\varphi \tag{10.9}$$

The transformed matrix is

$$RAR^{-1} = \begin{pmatrix} c & -s \\ s & c \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$
$$= \begin{pmatrix} c^2 a_{11} + s^2 a_{22} - 2csa_{12} & cs(a_{11} - a_{22}) + (c^2 - s^2)a_{12} \\ cs(a_{11} - a_{22}) + (c^2 - s^2)a_{12} & s^2 a_{11} + c^2 a_{22} + 2csa_{12} \end{pmatrix}. \tag{10.10}$$

It is diagonal if

$$0 = cs(a_{11} - a_{22}) + (c^2 - s^2)a_{12} = \frac{a_{11} - a_{22}}{2} \sin(2\varphi) + a_{12} \cos(2\varphi) \tag{10.11}$$

or

$$\tan(2\varphi) = \frac{2a_{12}}{a_{22} - a_{11}}. \tag{10.12}$$

Calculation of φ is not necessary since only its cosine and sine appear in (10.10). From [113]

$$\frac{1 - t^2}{t} = \frac{c^2 - s^2}{2cs} = \cot(2\varphi) = \frac{a_{22} - a_{11}}{2a_{12}} \tag{10.13}$$

we see that t is a root of

$$t^2 + \frac{a_{22} - a_{11}}{a_{12}} t - 1 = 0 \tag{10.14}$$

hence

$$t = -\frac{a_{22} - a_{11}}{2a_{12}} \pm \sqrt{1 + \left(\frac{a_{22} - a_{11}}{2a_{12}}\right)^2} = \frac{1}{\frac{a_{22} - a_{11}}{2a_{12}} \pm \sqrt{1 + \left(\frac{a_{22} - a_{11}}{2a_{12}}\right)^2}}. \tag{10.15}$$

For reasons of convergence [113] the solution with smaller magnitude is chosen which can be written as

$$t = \frac{\text{sign}\left(\frac{a_{22} - a_{11}}{2a_{12}}\right)}{\left|\frac{a_{22} - a_{11}}{2a_{12}}\right| + \sqrt{1 + \left(\frac{a_{22} - a_{11}}{2a_{12}}\right)^2}}. \tag{10.16}$$

Again for reasons of convergence the smaller solution φ is preferred and therefore we take

$$c = \frac{1}{\sqrt{1+t^2}} \quad s = \frac{t}{\sqrt{1+t^2}}. \tag{10.17}$$

The diagonal elements of the transformed matrix are

$$\tilde{a}_{11} = c^2 a_{11} + s^2 a_{22} - 2csa_{12} \tag{10.18}$$

$$\tilde{a}_{22} = s^2 a_{11} + c^2 a_{22} + 2csa_{12}. \tag{10.19}$$

The trace of the matrix is invariant

$$\tilde{a}_{11} + \tilde{a}_{22} = a_{11} + a_{22} \tag{10.20}$$

whereas the difference of the diagonal elements is

$$\begin{aligned} \tilde{a}_{11} - \tilde{a}_{22} &= (c^2 - s^2)(a_{11} - a_{22}) - 4csa_{12} = \frac{1-t^2}{1+t^2}(a_{11} - a_{22}) - 4\frac{a_{12}t}{1+t^2} \\ &= (a_{11} - a_{22}) + (-a_{12}\frac{1-t^2}{t})\frac{-2t^2}{1+t^2} - 4\frac{a_{12}t}{1+t^2} = (a_{11} - a_{22}) - 2ta_{12} \end{aligned} \tag{10.21}$$

and the transformed matrix has the simple form

$$\begin{pmatrix} a_{11} - a_{12}t & \\ & a_{22} + a_{12}t \end{pmatrix}. \tag{10.22}$$

For larger dimension $N > 2$ the Jacobi method uses the following algorithm:

(1) look for the dominant non-diagonal element $\max_{i \neq j} |a_{ij}|$

(2) Perform a rotation in the (ij)-plane to cancel the element $\tilde{a}_{ij}$ of the transformed matrix $\tilde{A} = R^{(ij)} \cdot A \cdot R^{(ij)-1}$. The corresponding rotation matrix has the form

$$R^{(ij)} = \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & c & & s & & \\ & & & \ddots & & & \\ & & -s & & c & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{pmatrix}. \tag{10.23}$$

(3) repeat (1–2) until convergence (if possible).

The sequence of Jacobi rotations gives the over all transformation

$$RAR^{-1} = \cdots R_2 R_1 A R_1^{-1} R_2^{-1} \cdots = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix}. \tag{10.24}$$

Hence

$$AR^{-1} = R^{-1} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_N \end{pmatrix} \tag{10.25}$$

and the column vectors of $R^{-1} = (\mathbf{v}_1, \mathbf{v}_2 \cdots \mathbf{v}_N)$ are the eigenvectors of A:

$$A\,(\mathbf{v}_1, \mathbf{v}_2 \cdots \mathbf{v}_N) = (\lambda_1 \mathbf{v}_1, \lambda_2 \mathbf{v}_2, \cdots \lambda_N \mathbf{v}_N)\,. \tag{10.26}$$

10.3 Tridiagonal Matrices

A tridiagonal matrix has nonzero elements only in the main diagonal and the first diagonal above and below. Many algorithms simplify significantly when applied to tridiagonal matrices.

10.3.1 Characteristic Polynomial of a Tridiagonal Matrix

The characteristic polynomial of a tridiagonal matrix

$$P_A(\lambda) = \det \begin{vmatrix} a_{11} - \lambda & a_{12} & & \\ a_{21} & a_{22} - \lambda & & \\ & & \ddots & a_{N-1N} \\ & & a_{NN-1} & a_{NN} - \lambda \end{vmatrix} \tag{10.27}$$

can be calculated recursively:

$$\begin{aligned} P_0 &= 1 \\ P_1(\lambda) &= a_{11} - \lambda \\ P_2(\lambda) &= (a_{22} - \lambda) P_1(\lambda) - a_{12} a_{21} \\ &\vdots \\ P_N(\lambda) &= (a_{NN} - \lambda) P_{N-1}(\lambda) - a_{N,N-1} a_{N-1,N} P_{N-2}(\lambda). \end{aligned} \tag{10.28}$$

10.3.2 Special Tridiagonal Matrices

Certain classes of tridiagonal matrices can be diagonalized exactly [114–116].

10.3.2.1 Discretized Second Derivatives

Discretization of a second derivative involves, under Dirichlet boundary conditions $f(x_0) = f(x_{N+1}) = 0$, the differentiation matrix (Sect. 20.2)

$$M = \begin{pmatrix} -2 & 1 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 1 & -2 \end{pmatrix}. \tag{10.29}$$

Its eigenvectors have the form

$$\boldsymbol{f} = \begin{pmatrix} f_1 \\ \vdots \\ f_n \\ \vdots \\ f_N \end{pmatrix} = \begin{pmatrix} \sin k \\ \vdots \\ \sin(nk) \\ \vdots \\ \sin(Nk) \end{pmatrix}. \tag{10.30}$$

This can be seen by inserting (10.30) into the n-th line of the eigenvalue (10.31)

$$M\boldsymbol{f} = \lambda \boldsymbol{f} \tag{10.31}$$

$$\begin{aligned} (M\boldsymbol{f})_n &= (\sin((n-1)k) + \sin((n+1)k) - 2\sin(nk)) \\ &= 2\sin(nk)\,(\cos(k) - 1) = \lambda\,(\boldsymbol{f})_n \end{aligned} \tag{10.32}$$

with the eigenvalue

$$\lambda = 2\,(\cos k - 1) = -4\sin^2\left(\frac{k}{2}\right). \tag{10.33}$$

The first line of the eigenvalue (10.31) reads

$$\begin{aligned} (M\boldsymbol{f})_1 &= (-2\sin(k) + \sin(2k)) \\ &= 2\sin(k)(\cos(k) - 1) = \lambda(\boldsymbol{f})_n \end{aligned} \tag{10.34}$$

and from the last line we have

$$\begin{aligned}(M\boldsymbol{f})_N &= (-2\sin(Nk) + \sin([N-1]k)) \\ &= \lambda(\boldsymbol{f})_N = 2(\cos(k) - 1)\sin(Nk)\end{aligned} \tag{10.35}$$

which holds if

$$\sin((N-1)k) = 2\sin(Nk)\cos(k). \tag{10.36}$$

This simplifies to

$$\begin{aligned}&\sin(Nk)\cos(k) - \cos(Nk)\sin(k) = 2\sin(Nk)\cos(k) \\ &\sin(Nk)\cos(k) + \cos(Nk)\sin(k) = 0 \\ &\sin((N+1)k) = 0.\end{aligned} \tag{10.37}$$

Hence the possible values of k are

$$k = \frac{\pi}{(N+1)} l \text{ with } l = 1, 2, \cdots N \tag{10.38}$$

and the eigenvectors are explicitly (Fig. 10.1)

$$\boldsymbol{f} = \begin{pmatrix} \sin\left(\frac{\pi}{N+1} l\right) \\ \vdots \\ \sin\left(\frac{\pi}{N+1} l\, n\right) \\ \vdots \\ \sin\left(\frac{\pi}{N+1} l\, N\right) \end{pmatrix}. \tag{10.39}$$

For Neumann boundary conditions $\frac{\partial f}{\partial x}(x_1) = \frac{\partial f}{\partial x}(x_N) = 0$ the matrix is slightly different (Sect. 20.2)

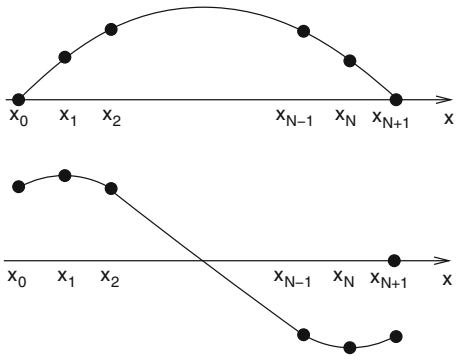

Fig. 10.1 (Lowest eigenvector) ***Top*** for fixed boundaries $f_n = \sin(nk)$ which is zero at the additional points x_0, x_{N+1}. ***Bottom*** for open boundaries $f_n = \cos((n-1)k)$ with horizontal tangent at x_1, x_N due to the boundary conditions $f_2 = f_0, f_{N-1} = f_{N+1}$

$$M = \begin{pmatrix} -2 & 2 & & & \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ & & & 2 & -2 \end{pmatrix}. \tag{10.40}$$

Its eigenvalues are also given by the expression (10.33). To obtain the eigenvectors, we try a more general ansatz with a phase shift

$$\boldsymbol{f} = \begin{pmatrix} \sin \Phi_1 \\ \vdots \\ \sin(\Phi_1 + (n-1)k) \\ \vdots \\ \sin(\Phi_1 + (N-1)k) \end{pmatrix}. \tag{10.41}$$

Obviously

$$\begin{aligned} &\sin(\Phi_1 + (n-1)k - k) + \sin(\Phi_1 + (n-1)k + k) - 2\sin(\Phi_1 + (n-1)k) \\ &= 2\,(\cos k - 1)\sin(\Phi_1 + (n-1)k). \end{aligned} \tag{10.42}$$

The first and last lines of the eigenvalue equation give

$$\begin{aligned} 0 &= -2\sin(\Phi_1) + 2\sin(\Phi_1 + k) - 2(\cos k - 1)\sin(\Phi_1) \\ &= 2\cos\Phi_1 \sin k \end{aligned} \tag{10.43}$$

and

$$\begin{aligned} 0 &= -2\sin(\Phi_1 + (N-1)k) + 2\sin(\Phi_1 + (N-1)k - k) \\ &\quad - 2(\cos k - 1)\sin(\Phi_1 + (N-1)k) = 2\cos(\Phi_1 + (N-1)k)\sin k \end{aligned} \tag{10.44}$$

which is solved by

$$\Phi_1 = \frac{\pi}{2} \quad k = \frac{\pi}{N-1} l, \quad l = 1, 2 \ldots N \tag{10.45}$$

hence finally the eigenvector is (Fig. 10.1)

$$\boldsymbol{f} = \begin{pmatrix} 1 \\ \vdots \\ \cos\left(\frac{n-1}{N-1}\pi l\right) \\ \vdots \\ (-1)^l \end{pmatrix}. \tag{10.46}$$

Even simpler is the case of the corresponding cyclic tridiagonal matrix

$$M = \begin{pmatrix} -2 & 1 & & & 1 \\ 1 & -2 & 1 & & \\ & \ddots & \ddots & \ddots & \\ & & 1 & -2 & 1 \\ 1 & & & 1 & -2 \end{pmatrix} \tag{10.47}$$

which has eigenvectors

$$\boldsymbol{f} = \begin{pmatrix} \mathrm{e}^{\mathrm{i}k} \\ \vdots \\ \mathrm{e}^{\mathrm{i}nk} \\ \vdots \\ \mathrm{e}^{\mathrm{i}Nk} \end{pmatrix} \tag{10.48}$$

and eigenvalues

$$\lambda = -2 + \mathrm{e}^{-\mathrm{i}k} + \mathrm{e}^{\mathrm{i}k} = 2\left(\cos(k) - 1\right) = -4\sin^2\left(\frac{k}{2}\right) \tag{10.49}$$

where the possible $k-$ values again follow from the first and last line

$$-2\mathrm{e}^{\mathrm{i}k} + \mathrm{e}^{\mathrm{i}2k} + \mathrm{e}^{\mathrm{i}Nk} = \left(-2 + \mathrm{e}^{-\mathrm{i}k} + \mathrm{e}^{\mathrm{i}k}\right)\mathrm{e}^{\mathrm{i}k} \tag{10.50}$$

$$\mathrm{e}^{\mathrm{i}k} + \mathrm{e}^{\mathrm{i}(N-1)k} - 2\mathrm{e}^{\mathrm{i}Nk} = \left(-2 + \mathrm{e}^{-\mathrm{i}k} + \mathrm{e}^{\mathrm{i}k}\right)\mathrm{e}^{\mathrm{i}Nk} \tag{10.51}$$

which both lead to

$$\mathrm{e}^{\mathrm{i}Nk} = 1 \tag{10.52}$$

$$k = \frac{2\pi}{N}l, \quad l = 0, 1, \cdots N-1. \tag{10.53}$$

10.3.2.2 Discretized First Derivatives

Using symmetric differences to discretize a first derivative in one dimension leads to the matrix[2]

[2]This matrix is skew symmetric, hence $\mathrm{i}T$ is Hermitian and has real eigenvalues $\mathrm{i}\lambda$.

$$D = \begin{pmatrix} & 1 & & & \\ -1 & & 1 & & \\ & \ddots & & \ddots & \\ & & \ddots & & \ddots \\ & & & -1 & & 1 \\ & & & & -1 & \end{pmatrix}. \tag{10.54}$$

The characteristic polynomial of the Hermitian matrix iD is given by the recursion Sect. 10.3.1

$$\begin{aligned} P_0 &= 1 \\ P_1 &= -\lambda \\ &\vdots \\ P_N &= -\lambda P_{N-1} - P_{N-2} \end{aligned} \tag{10.55}$$

which after the substitution $x = -\lambda/2$ is exactly the recursion for the Chebyshev polynomial of the second kind $U_N(x)$. Hence the eigenvalues of D are given by the roots x_k of $U_N(x)$ as

$$\lambda_D = 2\mathrm{i}x_k = 2\mathrm{i}\cos\left(\frac{k\pi}{N+1}\right) \quad k = 1, 2 \ldots N. \tag{10.56}$$

The eigenvalues of the corresponding cyclic tridiagonal matrix

$$D = \begin{pmatrix} & 1 & & & & -1 \\ -1 & & 1 & & & \\ & \ddots & & \ddots & & \\ & & \ddots & & \ddots & \\ & & & -1 & & 1 \\ 1 & & & & -1 & \end{pmatrix} \tag{10.57}$$

are easy to find. Inserting the ansatz for the eigenvector

$$\begin{pmatrix} \exp \mathrm{i}k \\ \vdots \\ \exp \mathrm{i}Nk \end{pmatrix} \tag{10.58}$$

we find the eigenvalues

$$\mathrm{e}^{\mathrm{i}(m+1)k} - \mathrm{e}^{\mathrm{i}(m-1)k} = \lambda \mathrm{e}^{\mathrm{i}mk} \tag{10.59}$$

$$\lambda = 2\mathrm{i}\sin k \tag{10.60}$$

and from the first and last equation

$$1 = \mathrm{e}^{\mathrm{i}Nk} \tag{10.61}$$

$$\mathrm{e}^{\mathrm{i}k} = \mathrm{e}^{\mathrm{i}(N+1)k} \tag{10.62}$$

the possible k-values

$$k = \frac{2\pi}{N} l, \quad l = 0, 1, \ldots N-1. \tag{10.63}$$

10.4 Reduction to a Tridiagonal Matrix

Eigenproblem algorithms work especially efficient if the matrix is first transformed to tridiagonal form (for real symmetric matrices, upper Hessian form for real non-symmetric matrices) which can be achieved by a series of Householder transformations (5.56)

$$A' = PAP \quad \text{with} \quad P = P^T = 1 - 2\frac{\boldsymbol{u}\boldsymbol{u}^T}{|\boldsymbol{u}|^2}. \tag{10.64}$$

The following orthogonal transformation P_1 brings the first row and column to tridiagonal form. We divide the matrix A according to

$$A = \begin{pmatrix} a_{11} & \boldsymbol{\alpha}^T \\ \boldsymbol{\alpha} & A_{rest} \end{pmatrix} \tag{10.65}$$

with the $(N-1)$- dimensional vector

$$\boldsymbol{\alpha} = \begin{pmatrix} a_{12} \\ \vdots \\ a_{1n} \end{pmatrix}.$$

Now let

$$\boldsymbol{u} = \begin{pmatrix} 0 \\ a_{12} + \lambda \\ \vdots \\ a_{1N} \end{pmatrix} = \begin{pmatrix} 0 \\ \boldsymbol{\alpha} \end{pmatrix} + \lambda \boldsymbol{e}^{(2)} \quad \text{with} \quad \boldsymbol{e}^{(2)} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{10.66}$$

Then

$$|\boldsymbol{u}|^2 = |\boldsymbol{\alpha}|^2 + \lambda^2 + 2\lambda a_{12} \tag{10.67}$$

and

$$\boldsymbol{u}^T \begin{pmatrix} a_{11} \\ \boldsymbol{\alpha} \end{pmatrix} = |\alpha|^2 + \lambda a_{12}. \tag{10.68}$$

The first row of A is transformed by multiplication with P_1 according to

$$P_1 \begin{pmatrix} a_{11} \\ \boldsymbol{\alpha} \end{pmatrix} = \begin{pmatrix} a_{11} \\ \boldsymbol{\alpha} \end{pmatrix} - 2\frac{|\alpha|^2 + \lambda a_{12}}{|\boldsymbol{\alpha}|^2 + \lambda^2 + 2\lambda a_{12}} \left[\begin{pmatrix} 0 \\ \boldsymbol{\alpha} \end{pmatrix} + \lambda \boldsymbol{e}^{(2)} \right]. \tag{10.69}$$

The elements number $3 \ldots N$ are eliminated if we choose[3]

$$\lambda = \pm|\alpha| \tag{10.70}$$

because then

$$2\frac{|\alpha|^2 + \lambda a_{12}}{|\boldsymbol{\alpha}|^2 + \lambda^2 + 2\lambda a_{12}} = 2\frac{|\alpha|^2 \pm |\alpha| a_{12}}{|\alpha|^2 + |\alpha|^2 \pm 2|\alpha| a_{12}} = 1 \tag{10.71}$$

and

$$P_1 \begin{pmatrix} a_{11} \\ \boldsymbol{\alpha} \end{pmatrix} = \begin{pmatrix} a_{11} \\ \boldsymbol{\alpha} \end{pmatrix} - \begin{pmatrix} 0 \\ \boldsymbol{\alpha} \end{pmatrix} - \lambda \boldsymbol{e}^{(2)} = \begin{pmatrix} a_{11} \\ \mp|\alpha| \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{10.72}$$

Finally we have

$$A^{(2)} = P_1 A P_1 = \begin{pmatrix} a_{11} & a_{12}^{(2)} & 0 & \cdots & 0 \\ a_{12}^{(2)} & a_{22}^{(2)} & a_{23}^{(2)} & \cdots & a_{2N}^{(2)} \\ 0 & a_{23}^{(2)} & \ddots & & a_{3N}^{(2)} \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & a_{2N}^{(2)} & a_{3N}^{(2)} & \cdots & a_{NN}^{(2)} \end{pmatrix} \tag{10.73}$$

as desired.

[3] To avoid numerical extinction we choose the sign to be that of a_{12}.

For the next step we choose

$$\boldsymbol{\alpha} = \begin{pmatrix} a_{22}^{(2)} \\ \vdots \\ a_{2N}^{(2)} \end{pmatrix}, \quad \boldsymbol{u} = \begin{pmatrix} 0 \\ 0 \\ \alpha \end{pmatrix} \pm |\alpha| \boldsymbol{e}^{(3)} \tag{10.74}$$

to eliminate the elements $a_{24} \ldots a_{2N}$. Note that P_2 does not change the first row and column of $A^{(2)}$ and therefore

$$A^{(3)} = P_2 A^{(2)} P_2 = \begin{pmatrix} a_{11} & a_{12}^{(2)} & 0 & \cdots\cdots & 0 \\ a_{12}^{(2)} & a_{22}^{(2)} & a_{23}^{(3)} & 0 \cdots & 0 \\ 0 & a_{23}^{(3)} & a_{33}^{(3)} & \cdots\cdots & a_{3N}^{(3)} \\ \vdots & 0 & \vdots & & \vdots \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{3N}^{(3)} & \cdots\cdots & a_{NN}^{(3)} \end{pmatrix}. \tag{10.75}$$

After $N-1$ transformations finally a tridiagonal matrix is obtained.

10.5 The Power Iteration Method

A real symmetric $N \times N$ matrix with (orthonormal) eigenvectors and eigenvalues[4]

$$A\boldsymbol{u}_i = \lambda_i \boldsymbol{u}_i \tag{10.76}$$

can be expanded as

$$A = \sum_i \boldsymbol{u}_i \lambda_i \boldsymbol{u}_i^T. \tag{10.77}$$

The sequence of powers

$$A^n = \sum_i \boldsymbol{u}_i \lambda_i^n \boldsymbol{u}_i^T \tag{10.78}$$

converges to

$$A^n \to \boldsymbol{u}_{max} \lambda_{max}^n \boldsymbol{u}_{max}^T$$

[4]We do not consider degenerate eigenvalues explicitly here.

where[5]

$$|\lambda_{max}| = \max. \tag{10.79}$$

Hence for any initial vector $\boldsymbol{v}_1$ (which is arbitrary but not perpendicular to $\boldsymbol{u}_{max}$) the sequence

$$\boldsymbol{v}_{n+1} = A\boldsymbol{v}_n \tag{10.80}$$

converges to a multiple of $\boldsymbol{u}_{max}$. To obtain all eigenvectors simultaneously, we could use a set of independent start vectors, e.g. the N unit vectors and iterate simultaneously for all of them

$$(\boldsymbol{v}_1^{(1)}, \ldots \boldsymbol{v}_1^{(N)}) = (\boldsymbol{e}_1, \ldots \boldsymbol{e}_N) = \begin{pmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{pmatrix} \tag{10.81}$$

$$(\boldsymbol{v}_2^{(1)}, \ldots \boldsymbol{v}_2^{(N)}) = A(\boldsymbol{v}_1^{(1)}, \ldots \boldsymbol{v}_1^{(N)}) = A. \tag{10.82}$$

Most probably, all column vectors of A then converge to multiples of the same eigenvector. To assure linear independence, an orthogonalization step has to follow each iteration. This can be done (QR decomposition, Sect. 5.2) by decomposing the matrix into the product of an upper triangular[6] matrix R and an orthogonal matrix $Q^T = Q^{-1}$ (Sect. 5.2)

$$A = QR. \tag{10.83}$$

For symmetric tridiagonal matrices this factorization can be efficiently realized by multiplication with a sequence of Givens rotation matrices which eliminate the off-diagonal elements in the lower part one by one[7]

$$Q = R_{\alpha_{N-1}}^{(N-1,N)} \ldots R_{\alpha_2}^{(2,3)} R_{\alpha_1}^{(1,2)} \tag{10.84}$$

beginning with

$$R_{\alpha_1}^{(1,2)} A = \begin{pmatrix} c & s & & & \\ -s & c & & & \\ & & 1 & & \\ & & & \ddots & \\ & & & & 1 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & & & & \\ a_{12} & a_{22} & a_{23} & & & \\ & \ddots & \ddots & \ddots & & \\ & & a_{N-2,N-1} & a_{N-1,N-1} & a_{N-1,N} & \\ & & & a_{N-1,N} & a_{N,N} & \end{pmatrix}$$

[5] For simplicity we do not consider eigenvalues which are different but have the same absolute value.

[6] The equivalent QL method uses a lower triangular matrix.

[7] This is quite different from the Jacobi method since it is not an orthogonal transformation.

$$= \begin{pmatrix} ca_{11} + sa_{12} & ca_{12} + sa_{22} & sa_{23} & & \\ 0 & -sa_{12} + ca_{22} & ca_{23} & & \\ & a_{34} & a_{33} & a_{34} & \\ & & \ddots & \ddots & \ddots \\ & & & a_{N-1,N} & a_{N,N} \end{pmatrix} \tag{10.85}$$

where the rotation angle α_1 has to be chosen such that

$$\tan \alpha_1 = \frac{s}{c} = \frac{a_{12}}{a_{11}}. \tag{10.86}$$

Finally, this leads to a method known as orthogonal simultaneous power iteration

$$W^{(1)} = A = Q^{(1)} R^{(1)} \tag{10.87}$$

$$W^{(n+1)} = AQ^{(n)} \tag{10.88}$$

$$Q^{(n+1)} R^{(n+1)} = W^{(n+1)}. \tag{10.89}$$

This method calculates a sequence of orthogonal matrices $Q^{(n)}$ which converge to a set of independent eigenvectors. Moreover, from (10.88) and (10.89)

$$A = W^{(n+1)} Q^{(n)T} = Q^{(n+1)} R^{(n+1)} Q^{(n)T} \tag{10.90}$$

and therefore powers of A are given by

$$\begin{aligned} A^n &= \left(Q^{(n)} R^{(n)} Q^{(n-1)T}\right) \left(Q^{(n-1)} R^{(n-1)} Q^{(n-2)T}\right) \ldots \left(Q^{(2)} R^{(2)} Q^{(1)T}\right) \left(Q^{(1)} R^{(1)}\right) \\ &= Q^{(n)} R^{(n)} R^{(n-1)} \ldots R^{(1)}. \end{aligned} \tag{10.91}$$

The product of two upper triangular matrices is upper triangular again which can be seen from

$$(R^{(m)} R^{(n)})_{i,k} = \sum_{j:\, i \le j \le k} R^{(m)}_{i,j} R^{(n)}_{jk} = 0 \text{ if } i > k. \tag{10.92}$$

Therefore the QR decomposition of A^n is

$$A^n = Q^{(n)} \overline{R}^{(n)} \tag{10.93}$$

with

$$\overline{R}^{(n)} = R^{(n)} \ldots R^{(1)}. \tag{10.94}$$

To obtain other than the dominant eigenvalues, the inverse power iteration method with shift is useful. Consider the matrix

$$\tilde{A} = (A - \sigma)^{-1} \tag{10.95}$$

where μ is not an eigenvalue $\lambda_i$of A. Obviously it has the same eigenvectors as A and eigenvalues given by

$$\tilde{A}\boldsymbol{u}_i = \tilde{\lambda}_i \boldsymbol{u}_i = \frac{1}{\lambda_i - \sigma}\boldsymbol{u}_i. \tag{10.96}$$

Hence, if σ is close to λ_i, the power iteration method will converge to a multiple of $\boldsymbol{u}_i$. For practical calculations, an equivalent formulation of the power iteration method is used which is known as the QR (or QL) method.

10.6 The QR Algorithm

The QR algorithm [117] is an iterative algorithm. It uses a series of orthogonal transformations which conserve the eigenvalues. Starting from the decomposition of A

$$A = Q_1 R_1 \tag{10.97}$$

$$A_2 = R_1 Q_1 = Q_1^T A Q_1 \tag{10.98}$$

we iterate

$$\vdots$$

$$A_n = Q_n R_n \tag{10.99}$$

$$A_{n+1} = R_n Q_n = Q_n^T A_n Q_n. \tag{10.100}$$

From (10.99) and (10.100)

$$Q_{n+1} R_{n+1} = R_n Q_n$$

and the n-th power of A is

$$\begin{aligned} A^n &= AA \ldots A = Q_1 R_1 Q_1 R_1 \ldots Q_1 R_1 = Q_1 (Q_2 R_2 \ldots Q_2 R_2) R_1 \\ &= Q_1 Q_2 (Q_3 R_3 \ldots Q_3 R_3) R_2 R_1 \cdots = \overline{Q}_n \overline{R}_n \end{aligned} \tag{10.101}$$

$$\overline{Q}_n = Q_1 \ldots Q_n \quad \overline{R}_n = R_n \ldots R_1. \tag{10.102}$$

But since QR decomposition is unique, comparison of (10.93) and (10.101) shows

$$\overline{Q}_n = Q^{(n)} \quad \overline{R}_n = \overline{R}^{(n)} \tag{10.103}$$

i.e. the column vectors of $\overline{Q}_n$ converge to a set of eigenvectors and the transformed matrix

$$A_{n+1} = Q_n^T A_n Q_n = Q_n^T Q_{n-1}^T A_{n-1} Q_{n-1} Q_n = \cdots = \overline{Q}_n^T A \overline{Q}_n \tag{10.104}$$

converges to a diagonal matrix. Now consider the inverse power

$$A^{-n} = \overline{R}_n^{-1} \overline{Q}_n^T. \tag{10.105}$$

The inverse of a symmetric matrix is also symmetric and

$$A^{-n} = \overline{Q}_n \left(\overline{R}_n^{-1}\right)^T \tag{10.106}$$

shows, that the QR algorithm uses the same orthogonal transformations as ordinary and also inverse power iteration. The inverse of an upper triangular matrix is also upper triangular but the transpose is lower triangular. Therefore we modify (10.106) by multiplying with a permutation matrix

$$P = \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix} \quad P^2 = 1 \tag{10.107}$$

from the right side, which reverses the order of the columns and

$$A^{-n} P = \overline{Q}_n P P \left(\overline{R}_n^{-1}\right)^T P = \tilde{Q}\,\tilde{R} \tag{10.108}$$

is the QR decomposition of $A^{-n}P$. This shows the close relationship between the QR algorithm[8] and the inverse power iteration method.

To improve convergence, a shift σ is introduced and the QR factorization applied to $A_n - \sigma$. The modified iteration then reads

$$A_n - \sigma = Q_n R_n \tag{10.109}$$

$$A_{n+1} = R_n Q_n + \sigma = Q_n^T (A_n - \sigma) Q_n + \sigma = Q_n^T A_n Q_n. \tag{10.110}$$

[8] Or the equivalent QL algorithm.

Symmetry and tridiagonal form are conserved by this algorithm. The simplest choice for the shift is to take a diagonal element $\sigma_m = a_{m,m} = (A_n)_{m,m}$ corresponding to the Rayleigh quotient method. An even more robust and very popular choice [113] is Wilkinson's shift

$$\sigma_m = a_{m,m} + \delta - \text{sign}(\delta)\sqrt{\delta^2 + a_{m,m-1}^2} \quad \delta = \frac{a_{m-1,m-1} - a_{m,m}}{2} \tag{10.111}$$

which is that eigenvalue of the matrix $\begin{pmatrix} a_{m-1,m-1} & a_{m-1,m} \\ a_{m-1,m} & a_{m,m} \end{pmatrix}$ which is closer to $a_{n,n}$. The calculation starts with σ_N and iterates until the off-diagonal element $a_{N-1,N}$ becomes sufficiently small.[9] Then the transformed matrix has the form

$$\begin{pmatrix} a_{11} & a_{12} & & & & \\ a_{12} & a_{22} & \ddots & & & \\ & \ddots & \ddots & a_{N-2,N-1} & & \\ & & a_{N-2,N-1} & a_{N-1,N-1} & 0 & \\ & & & 0 & a_{NN} & \end{pmatrix}. \tag{10.112}$$

Now the last column and row can be discarded (deflation method) and the next iteration performed with the shift σ_{N-1} on a tridiagonal matrix of dimension $N-1$. This procedure has to be repeated N times to obtain all eigenvalues. Convergence is usually cubic (or at least quadratic if there are degenerate eigenvalues).

10.7 Hermitian Matrices

In quantum mechanics often Hermitian matrices have to be diagonalized (which have real valued eigenvalues). To avoid complex arithmetics, an Hermitian eigenproblem can be replaced by a symmetric real valued problem of double dimension by introducing

$$B = \Re(A) \quad C = \Im(A) \quad \boldsymbol{x} = \boldsymbol{u} + \mathrm{i}\boldsymbol{v} \tag{10.113}$$

where, for Hermitian A

$$A = B + \mathrm{i}C = A^H = B^T - \mathrm{i}C^T \tag{10.114}$$

hence

$$B = B^T \quad C = -C^T \tag{10.115}$$

[9] For the QL method, it is numerically more efficient to start at the upper left corner of the matrix.

and the eigenvalue problem can be rewritten as

$$0 = (A\boldsymbol{x} - \lambda\boldsymbol{x}) = (B\boldsymbol{u} - C\boldsymbol{v} - \lambda\boldsymbol{u}) + \mathrm{i}(B\boldsymbol{v} + C\boldsymbol{u} - \lambda\boldsymbol{v}) \tag{10.116}$$

or finally in the real symmetric form

$$\begin{pmatrix} B & C^T \\ C & B \end{pmatrix} \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} = \lambda \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix}. \tag{10.117}$$

Each eigenvalue of the N-dimensional Hermitian problem corresponds to two eigenvectors of the 2N-dimensional problem since for any solution of (10.117)

$$\begin{pmatrix} B & -C \\ C & B \end{pmatrix} \begin{pmatrix} -\boldsymbol{v} \\ \boldsymbol{u} \end{pmatrix} = \begin{pmatrix} -B\boldsymbol{v} - C\boldsymbol{u} \\ -C\boldsymbol{v} + B\boldsymbol{u} \end{pmatrix} = \lambda \begin{pmatrix} -\boldsymbol{v} \\ \boldsymbol{u} \end{pmatrix} \tag{10.118}$$

provides a different solution, while the complex vectors $\boldsymbol{u} + \mathrm{i}\boldsymbol{v}$ and $\mathrm{i}(\boldsymbol{u} + \mathrm{i}\boldsymbol{v}) = -\boldsymbol{v} + \mathrm{i}\boldsymbol{u}$ only differ by a phase factor.

10.8 Large Matrices

Many problems in computational physics involve very large matrices, for which standard methods are not applicable. It might be even difficult or impossible to keep the full matrix in memory. Here methods are used which only involve the product of the matrix with a vector which can be computed on the fly. Krylov methods are very similar to power iteration but diagonalize only the projection of the matrix onto a Krylov space of much smaller dimension $n \ll N$ which is constructed by multiplying a normalized start vector $\boldsymbol{q}_1$ repeatedly with A

$$K_n(A, \boldsymbol{q}_1) = \mathrm{span}\{\boldsymbol{q}_1, A\boldsymbol{q}_1, A^2\boldsymbol{q}_1, \ldots A^{n-1}\boldsymbol{q}_1\}. \tag{10.119}$$

We use the Arnoldi method (Sect. 5.6.5) to construct an orthonormalized basis of this space. For a symmetric matrix this simplifies to a three-term recursion also known as symmetric Lanczos algorithm [118]. Applying the Arnoldi method

$$h_{j,n} = (\boldsymbol{q}_j^T A \boldsymbol{q}_n) \quad j \leq n \tag{10.120}$$

$$\tilde{\boldsymbol{q}}_{n+1} = A\boldsymbol{q}_n - \sum_{j=1}^{n} h_{jn}\boldsymbol{q}_j \tag{10.121}$$

$$h_{n+1,n} = |\tilde{\boldsymbol{q}}_{n+1}| \quad \boldsymbol{q}_{n+1} = \frac{\tilde{\boldsymbol{q}}_{n+1}}{h_{n+1,n}} \tag{10.122}$$

to a symmetric matrix A, we find

$$h_{n+1,n}^2 = \boldsymbol{q}_n^T A^2 \boldsymbol{q}_n - 2\sum_j (\boldsymbol{q}_n^T A \boldsymbol{q}_j)(\boldsymbol{q}_j^T A \boldsymbol{q}_n) + \sum_{jj'} (\boldsymbol{q}_j^T A \boldsymbol{q}_n)(\boldsymbol{q}_{j'}^T A \boldsymbol{q}_n)\delta_{jj'}$$
$$= \boldsymbol{q}_n^T A \left[h_{n+1,n}\boldsymbol{q}_{n+1} + \sum_j h_{jn}\boldsymbol{q}_j \right] - \sum_j h_{jn}^2 = h_{n+1,n}h_{n,n+1} \tag{10.123}$$

hence

$$h_{n+1,n} = h_{n,n+1}. \tag{10.124}$$

Furthermore,

$$h_{n-2,n} = \boldsymbol{q}_{n-2}^T A \boldsymbol{q}_n = \boldsymbol{q}_{n-2}^T A \frac{1}{h_{n,n-1}} \left[A\boldsymbol{q}_{n-1} - \sum_{j=1}^{n-1} h_{jn-1}\boldsymbol{q}_j \right]$$
$$= \frac{1}{h_{n.n-1}} \left[-\sum_{j=1}^{n-2} h_{jn-1}h_{j,n-2} - h_{n-1,n-1}h_{n-2,n-1} + \boldsymbol{q}_{n-1}^T A \left(h_{n-1,n-2}\boldsymbol{q}_{n-1} + \sum_{j=1}^{n-2} h_{jn-2}\boldsymbol{q}_j \right) \right]$$
$$= \frac{1}{h_{n.n-1}} \left[-\sum_{j=1}^{n-2} h_{jn-1}h_{j,n-2} - h_{n-1,n-1}h_{n-2,n-1} + h_{n-1,n-1}h_{n-1,n-2} + \sum_{j=1}^{n-2} h_{jn-1}h_{jn-2} \right] = 0 \tag{10.125}$$

and similar for $s > 2$

$$h_{n-s,n} = \boldsymbol{q}_{n-s}^T A \boldsymbol{q}_n = \boldsymbol{q}_{n-s}^T A \frac{1}{h_{n,n-1}} \left[A\boldsymbol{q}_{n-1} - \sum_{j=1}^{n-1} h_{jn-1}\boldsymbol{q}_j \right]$$
$$= \frac{1}{h_{n.n-1}} \left[-\sum_{j=1}^{n-2} h_{jn-1}h_{j,n-s} - h_{n-1,n-1}h_{n-s,n-1} + \boldsymbol{q}_{n-1}^T A \left(h_{n-s+1,n-s}\boldsymbol{q}_{n-s+1} + \sum_{j=1}^{n-s} h_{jn-s}\boldsymbol{q}_j \right) \right]$$
$$= \frac{1}{h_{n.n-1}} \left[-\sum_{j=1}^{n-2} h_{jn-1}h_{j,n-s} - h_{n-1,n-1}h_{n-s,n-1} + h_{n-s+1,n-s}h_{n-1,n-s+1} + \sum_{j=1}^{n-s} h_{jn-1}h_{jn-s} \right]$$
$$= \frac{1}{h_{n,n-1}} \left[-\sum_{j=n-s+1}^{n-2} h_{j,n-1}h_{j,n-s} - h_{n-1,n-1}h_{n-s,n-1} + h_{n-s,n-s+1}h_{n-s+1,n-1} \right] \tag{10.126}$$

$$h_{n-s,n} = \frac{1}{h_{n,n-1}} \left[-\sum_{j=n-s+2}^{n-2} h_{j,n-1}h_{j,n-s} - h_{n-1,n-1}h_{n-s,n-1} \right]. \tag{10.127}$$

Starting from (10.125) for $s = 2$ we increment s repeatedly and find

$$h_{n-2,n} = h_{n-3,n} = \ldots h_{1,n} = 0 \tag{10.128}$$

since (10.127) only involves smaller values of s, for which (10.128) already has been shown. The Arnoldi decomposition produces an upper Hessenberg matrix Sect. 5.6.5,

$$U_n = (\boldsymbol{q}_1, \ldots \boldsymbol{q}_n) \tag{10.129}$$

$$AU_n = U_{n+1}H = (U_n, \boldsymbol{q}_{n+1}) \begin{pmatrix} H_n \\ h_{n+1,n}\boldsymbol{e}_n^T \end{pmatrix} = U_n H_n + h_{n+1}\boldsymbol{q}_{n+1}\boldsymbol{e}_n^T \tag{10.130}$$

which for symmetric A becomes tridiagonal

$$H = \begin{pmatrix} h_{11} & h_{12} & \ldots & h_{1n} \\ h_{21} & h_{22} & \ldots & h_{2n} \\ & h_{32} & \ddots & \\ & & \ddots & h_{nn} \\ & & & h_{n+1,n} \end{pmatrix} = \begin{pmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & b_{n-2} & a_{n-1} & b_{n-1} \\ & & & b_{n-1} & a_n \\ & & & & b_n \end{pmatrix} = \begin{pmatrix} T \\ b_n\boldsymbol{e}_n^T \end{pmatrix} \tag{10.131}$$

with a symmetric tridiagonal matrix T, which is the desired projection of A into the Krylov space K_n

$$\begin{aligned} U_n^T A U_n &= U_n^T U_{n+1} H = U_n^T \ (U_n, \boldsymbol{q}_{n+1}) \begin{pmatrix} T \\ b_n\boldsymbol{e}_n^T \end{pmatrix} \\ &= (E_n, 0) \begin{pmatrix} T \\ b_n\boldsymbol{e}_n^T \end{pmatrix} = T. \end{aligned} \tag{10.132}$$

For an eigenpair $(\lambda, \boldsymbol{v})$ of T

$$\begin{aligned} A(U_n\boldsymbol{v}) &= U_{n+1}H\boldsymbol{v} = \left(U_n, \boldsymbol{q}_{n+1}\right) \begin{pmatrix} T \\ b_n\boldsymbol{e}_n^T \end{pmatrix} \boldsymbol{v} \\ &= \left(U_n T + b_n\boldsymbol{q}_{n+1}\boldsymbol{e}_n^T\right) \boldsymbol{v} = \lambda(U_n\boldsymbol{v}) + b_n\boldsymbol{q}_{n+1}\boldsymbol{e}_n^T\boldsymbol{v}. \end{aligned} \tag{10.133}$$

Hence, an approximate eigenpair of A is given by the Ritz pair $(\lambda, U_n\boldsymbol{v})$ and the error can be estimated from the residual norm

$$\frac{|(A-\lambda)U_n\boldsymbol{v}|}{|\boldsymbol{v}|} = |b_n||\boldsymbol{e}_n^T\boldsymbol{v}|. \tag{10.134}$$

Due to numerical errors, orthogonality of the Lanczos vectors $\boldsymbol{q}_n$ can get lost and reorthogonalization is necessary [119, 120]. If this takes to much time or if memory limits do not allow to store enough Lanczos vectors, the procedure has to be restarted with a new initial vector which is usually taken as a linear combination of selected eigenvectors which have already been found [121, 122]. Furthermore, special care has to be taken to determine possible degeneracies of the eigenvalues.

10.9 Non-symmetric Matrices

Eigenvalue problems with non-symmetric matrices are more complicated. Left and right eigenvectors have to be distinguished and the eigenvalues can be complex valued even if the matrix is real. The QR method [117] is applicable also to a non-symmetric matrix but very expensive unless the matrix is first brought to upper triangular (instead of tridiagonal) form, which can be achieved by a series of similarity transformations with Householder reflections (5.2.2). The implicit QR method with double shift avoids complex arithmetics by treating pairs of complex conjugated eigenvalues simultaneously. For very large matrices the Arnoldi method brings a non-symmetric matrix to upper Hessenberg form, which provides the projection onto the Krylov space as an upper triangular matrix.

Problems

Problem 10.1 Computer Experiment: Disorder in a Tight-Binding Model

We consider a two-dimensional lattice of interacting particles. Pairs of nearest neighbors have an interaction V and the diagonal energies are chosen from a Gaussian distribution

$$P(E) = \frac{1}{\Delta\sqrt{2\pi}} e^{-E^2/2\Delta^2}. \tag{10.135}$$

The wave function of the system is given by a linear combination

$$\psi = \sum_{ij} C_{ij}\psi_{ij} \tag{10.136}$$

where on each particle (i, j) one basis function ψ_{ij} is located. The nonzero elements of the interaction matrix are given by

$$H(ij|ij) = E_{ij} \tag{10.137}$$

$$H(ij|i \pm 1, j) = H(ij|i, j \pm 1) = V. \tag{10.138}$$

The Matrix H is numerically diagonalized and the amplitudes C_{ij} of the lowest state are shown as circles located at the grid points. As a measure of the degree of localization the quantity

$$\sum_{ij} |C_{ij}|^4 \tag{10.139}$$

is evaluated. Explore the influence of coupling V and disorder Δ.