# Numerical Methods in Linear Algebra 

Dénes Sexty

University of Graz

2020-2021, winter term

UNIVERSITÄT GRAZ UNIVERSITY OF GRAZ

UNI

1 Examples: Bridge, multidim Newton's method, QM, vibrational eigenmodes
2 Linear equations
3 Gauss elimination, Pivoting
4 Numerical errors, condition number
5 LU decomposition, iterative improvement
б Householder reduction
7 iterative solution: Gauss-Seidel, Successive overrelaxation (SOR), Conjugate gradient (CG) and others
8 Eigenvalues Subspace methods, power iteration, Lánczos method, Krylov Schur

9 Eigenvalues transformation method: Jacobi, QR
(10) Diagonalization with Fourier transformation

11 Partial Differential Equations: Boundary value problems, Finite elements
I2 Partial Differential Equations: initial value problems. Courant condition, Crank-Nicholson method

$$
\begin{equation*}
\left(1+q+q^{2}+\ldots+q^{N-1}\right)(1-q)=1-q^{N} \Longrightarrow \sum_{k=0}^{N-1} q^{k}=\frac{1-q^{N}}{1-q} \tag{1}
\end{equation*}
$$

Let's use it for $q=\exp (i(2 \pi / N) j)$ with $j \in \mathbb{Z}$
if $j \neq N /$ with $I \in \mathbb{Z} \Longrightarrow q^{N}=1, q \neq 1 \Longrightarrow$ sum is zero if $j=N /$ with $I \in \mathbb{Z} \quad \Longrightarrow q=1$

$$
\begin{equation*}
\sum_{k=0}^{N-1} q^{k}=N \tag{2}
\end{equation*}
$$

So in summary

$$
\begin{equation*}
\frac{1}{N} \sum_{k=0}^{N-1} e^{i \frac{2 \pi}{N} k j}=\delta_{j \bmod N, 0} \tag{3}
\end{equation*}
$$

where $j \bmod N$ is the remainder of the division of $j$ with $N$

## Fourier Transformation

Let's define the $N \times N$ matrix $U$

$$
\begin{equation*}
U_{k j}=\frac{1}{\sqrt{N}} e^{-i \frac{2 \pi}{N} k j}, \quad j, k=0,1, \ldots, N-1 \tag{4}
\end{equation*}
$$

The matrix is unitary

$$
\begin{equation*}
\left(U U^{+}\right)_{k l}=\sum_{j=0}^{N-1} U_{k j} U_{j l}^{+}=\sum_{j=0}^{N-1} \frac{1}{N} e^{-i \frac{2 \pi}{N} k j} e^{i \frac{2 \pi}{N} j l}=\frac{1}{N} \sum_{j=0}^{N-1} e^{i \frac{2 \pi}{N} j(l-k)}=\delta_{\mid k} \tag{5}
\end{equation*}
$$

Similarly we see $U^{+} U=1$

Fourier transformation of a vector:

$$
\begin{equation*}
\tilde{\mathbf{v}}=U \mathbf{v}, \quad \Longrightarrow U^{+} \tilde{\mathbf{v}}=U^{+} U \mathbf{v}=\mathbf{v} \tag{6}
\end{equation*}
$$

Fourier Transformation of a matrix:

$$
\begin{equation*}
\tilde{M}=U M U^{+}, \quad U^{+} \tilde{M} U=M \tag{7}
\end{equation*}
$$

Unitary transformation $\Longrightarrow M$ and $\tilde{M}$ have the same eigenvalues.

$$
\begin{equation*}
M^{-1}=U^{+} \tilde{M}^{-1} U \tag{8}
\end{equation*}
$$

Proof:

$$
\begin{array}{r}
M M^{-1}=M U^{+} \tilde{M}^{-1} U=U^{+} \tilde{M} \underbrace{U U^{+}}_{=1} \tilde{M}^{-1} U=  \tag{9}\\
=U^{+} \underbrace{\tilde{M} \tilde{M}^{-1}}_{=1} U=U^{+} U=\mathbf{1}
\end{array}
$$

Can use Fourier transformation to

- Diagonalization
- Calculation of inverse
- convolution, etc.
- Data analysis, Data filtering, etc.


## Diagonalization with Fourier Transformation

Consider the $N \times N$ matrix

$$
\begin{equation*}
M_{j k}=3 \delta_{j k}-\delta_{j+1, k}-\delta_{k-1, k} \tag{10}
\end{equation*}
$$

$0 \leq j, k \leq N-1$, using periodic boundary conditions: $N \rightarrow 0,-1 \rightarrow N-1$

$$
M=\left(\begin{array}{ccccc}
3 & -1 & 0 & \ldots & -1  \tag{11}\\
-1 & 3 & -1 & \ldots & 0 \\
0 & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & & \ddots & \ddots & -1 \\
-1 & 0 & \ldots & -1 & 3
\end{array}\right)
$$

Next we calculate $\tilde{M}=U M U^{+}$

## Diagonalization with Fourier Transformation2

Next we calculate $\tilde{M}=U M U^{+}$

$$
\begin{array}{r}
\left(U M U^{+}\right)_{k l}=\frac{1}{N} \sum_{n, m} e^{-i \frac{2 \pi}{N} k n}\left(3 \delta_{n m}-\delta_{n+1, m}-\delta_{n-1, m}\right) e^{i \frac{2 \pi}{N} / m}=  \tag{12}\\
=\frac{1}{N} \sum_{n} e^{-i \frac{2 \pi}{N} k n}\left(3 e^{i \frac{2 \pi}{N} / n}-e^{i \frac{2 \pi}{N} /(n+1)}-e^{i \frac{2 \pi}{N} /(n-1)}\right) \\
=\frac{1}{N} \sum_{n} e^{-i \frac{2 \pi}{N} k n} e^{i \frac{2 \pi}{N} n l}\left(3-e^{i \frac{2 \pi}{N} l}-e^{-i \frac{2 \pi}{N} l}\right) \\
=\frac{1}{N} \sum_{n} e^{-i \frac{2 \pi}{N} n(l-k)}(3-2 \cos (2 \pi I / N)) \\
=\delta_{l k}(3-2 \cos (2 \pi I / N))
\end{array}
$$

$\Longrightarrow \tilde{M}$ is diagonal

$$
\begin{equation*}
\lambda_{I}=(3-2 \cos (2 \pi I / N)) \tag{13}
\end{equation*}
$$

## Diagonalization with Fourier Transformation3

We had $\tilde{M}=U M U^{+}=D$ diagonal

$$
\begin{equation*}
M U^{+}=U^{+} D \tag{14}
\end{equation*}
$$

The columns of $U^{+}$are the eigenvectors of $M$
$\mathbf{v}_{0}=\frac{1}{\sqrt{N}}\left(\begin{array}{c}1 \\ 1 \\ \vdots \\ 1\end{array}\right), \quad \mathbf{v}_{1}=\frac{1}{\sqrt{N}}\left(\begin{array}{c}e^{i \frac{2 \pi}{N} 0} \\ e^{i \frac{2 \pi}{N} 1} \\ \vdots \\ e^{i \frac{2 \pi}{N}(N-1)}\end{array}\right), \ldots, \mathbf{v}_{N-1}=\frac{1}{\sqrt{N}}\left(\begin{array}{c}\frac{i \frac{2 \pi}{N} 0}{i \frac{2 \pi}{N}(N-1)} \\ \vdots \\ e^{i \frac{2 \pi}{N}(N-1)(N-1)}\end{array}\right)$
We can also calculate the inverse:

$$
\begin{equation*}
M_{n m}^{-1}=U^{+} \tilde{M}^{-1} U=\frac{1}{N} \sum_{k} \frac{e^{i \frac{2 \pi}{N} k(n-m)}}{\lambda_{k}} \tag{16}
\end{equation*}
$$

Discretisation of a first derivative:

$$
\begin{equation*}
\frac{f(x+a)-f(x)}{a}=\frac{1}{a}\left(f(x)+a f^{\prime}(x)+\ldots-f(x)\right) \approx f^{\prime}(a) \tag{17}
\end{equation*}
$$

Second derivative one could discretise e.g.:

$$
\begin{array}{r}
\frac{f(x+a)-2 f(x)+f(x-a)}{a^{2}}=\frac{1}{a^{2}}\left(f(x)+a f^{\prime}(x)+\frac{a^{2}}{2} f^{\prime \prime}(x)+\ldots+\right.  \tag{18}\\
\left.+f(x)-a f^{\prime}(x)+\frac{a^{2}}{2} f^{\prime \prime}(x)+\ldots-2 f(x)\right) \approx f^{\prime \prime}(a)
\end{array}
$$

Let's look at our matrix again: $M_{j k}=3 \delta_{j k}-\delta_{j+1, k}-\delta_{k-1, k}$ Apply this to a "function" $f_{x}$

$$
\begin{equation*}
(M f)_{x}=3 f_{x}-f_{x-1}-f_{x+1}=f_{x}-\left(f_{x-1}-2 f_{x}+f_{x+1}\right) \tag{19}
\end{equation*}
$$

$\Longrightarrow$ In the continuum, this corresponds to the operator: $N=1-\partial_{x}^{2}$
$\Longrightarrow$ In the continuum, this corresponds to the operator: $N=1-\partial_{x}^{2}$ After Fouerier transformation, this goes into $N(k)=1+k^{2}$

On the lattice (discretised version of the operator):
The eigenvalues were given:

$$
\begin{equation*}
\lambda_{I}=3-2 \cos (2 \pi I / N)=1+2-2 \cos (2 \pi I / N) \tag{20}
\end{equation*}
$$

Using $\cos (2 x)=1-2 \sin ^{2}(x)$

$$
\lambda_{k}=1+2-2 \cos (2 \pi k / N)=1+2-2\left(1-2 \sin ^{2}(\pi k / N)\right)=1+(2 \sin (\pi k / N))^{2}
$$

This suggests to define lattice momentum squared

$$
\begin{equation*}
k_{L A T}^{2}=(2 \sin (\pi k / N))^{2}, \quad k=0, \ldots, N-1 \tag{21}
\end{equation*}
$$

## Partial Differential equations

Many problems in physics: heat equation (or diffusion), wave equation in elastic media, Maxwell equation, time dependent Schrödinger equation or Dirac equation, Euler and Navier-Stokes equation, Klein-Gordon equation, etc. Analytical solutions rarely exist. they are often linear. Numerical solution needed.

Classification: consider a linear PDE of $u(x, y)$ which is at most second order The most general form is:

$$
\begin{equation*}
A u_{x x}+2 B u_{x y}+C u_{y y}+D u_{x}+E u_{y}+F u+G=0 \tag{22}
\end{equation*}
$$

where $u_{x x}$ is the second partial derivative with respect to $x$, etc. The coefficients $A, B, \ldots$ might depend on $x$ and $y$
This is analogous to a second order form

$$
\begin{equation*}
A x^{2}+2 B x y+C y^{2}+\ldots=0 \tag{23}
\end{equation*}
$$

which we know for constant $A, B, C$ describes conic sections (ellipsoid, paraboloid, hyperboloid)

Classification of conic sections: what curve is described by the following eq?
$\left(\begin{array}{ll}x y\end{array}\right)\left(\begin{array}{ll}A & B \\ B & C\end{array}\right)\binom{x}{y}+D x+E y=A x^{2}+2 B x y+C y^{2}+D x+E y=$ const
If the determinant of the matrix is nonzero, we can make $D$ and $E$ vanish by a shift of the origin: $x=x^{\prime}-x_{0}, y=y^{\prime}-y_{0}$
Than we do an orthogonal rotation to make the matrix diagonal. Then our eq. is: $A^{\prime \prime} x^{\prime \prime 2}+C^{\prime \prime} y^{\prime \prime 2}=$ const.

If the determinant of the matrix $A C-B^{2}$ is positive than both eigenvalues $A^{\prime \prime}$ and $C^{\prime \prime}$ have the same sign $\Longrightarrow$ ellipse

If the determinant of the matrix $A C-B^{2}$ is negative than the eigenvalues $A^{\prime \prime}$ and $C^{\prime \prime}$ have the opposite sign $\Longrightarrow$ hyperbola

If the determinant is zero $A C-B^{2}$, we can't do the shift, and after the rotation we end up with $A^{\prime} x^{\prime 2}+D^{\prime} x^{\prime}+E y^{\prime}=$ const.
Now doing the shift leads to $A^{\prime} x^{\prime \prime 2}+E^{\prime} y^{\prime \prime}=0$ which describes a parabola
$B^{2}-A C$ is usually called the discriminant

## PDE classifications

- $B^{2}-A C<0$ : elliptic PDEs.

The solutions are smooth if the coefficients are smooth. e.g.: Laplace equation $\triangle f=0$ has analytic solutions in a region even if the values on the border are not smooth.
prototypical example: Poisson equation for $u$ electrostatic potential, $\rho(x, y)$ charge distribution

$$
\begin{equation*}
u_{x x}+u_{y y}=-\rho(x, y) \tag{24}
\end{equation*}
$$

- $B^{2}-A C=0:$ parabolic PDEs.

Can be transformed to the heat equation. Solutions smooth out as time is increased. heat eq. for $u$ temperature and $\kappa$ diffusion coefficient:

$$
\begin{equation*}
u_{t}=\kappa u_{x x} \tag{25}
\end{equation*}
$$

- $B^{2}-A C>0$ : hyperbolic PDEs:

Wave equations. Discontinuities in the initial conditions are retained for later times. Wave eq. with speed of light (or sound, etc.) $v$ is:

$$
\begin{equation*}
u_{t t}=c^{2} u_{x x} \tag{26}
\end{equation*}
$$

## PDE classification

- Boundary value problems (aka. static problem) e.g.: Poisson equation a quantity $u(x, y)$ (or its derivative or some other property) is given on the boundary of some region, $u(x, y)$ inside the region is to be calculated.

Usually some kind of iterative procedure is employed, one looks for efficient solutions.
After discretisation of linear PDEs one gets a system of linear equations, typically with a sparse matrix.

■ initial value problems (aka. time evolution problem) e.g: Heat equation, wave equation
$u(x, t)$ is known at initial times and its time evolution is to be calculated. $u(x, t>0)$ must also be given on the boundary of the region of interest (periodic boundary conditions are also possible)

Main concern here is to devise algorithms which give a stable solution.

## Boundary value problems

We start with the Poisson equation (or Laplace equation for $\rho(\mathbf{x})=0$.

$$
\begin{equation*}
\sum_{i=1}^{d} \frac{\partial^{2}}{\partial x_{i}^{2}} \Phi(x)=\triangle \Phi(x)=-\rho(\mathbf{x}) \tag{27}
\end{equation*}
$$

where we look for $\Phi(\mathbf{x})$ in some region $\Omega$, with $\left.\Phi(\mathbf{x})\right|_{\mathbf{x} \in \partial \Omega}$ given. (The boundary can be the union of a not connected parts)
We discretise space using a cubic lattice:

$$
\begin{equation*}
x_{i}=x_{i 0}+a n_{i} \tag{28}
\end{equation*}
$$

Which means:

$$
\begin{equation*}
\mathbf{x}=\mathbf{n} a \tag{29}
\end{equation*}
$$

where $\mathbf{n}$ is a vector of integers.
This also means:

$$
\begin{equation*}
\Phi(\mathbf{x}) \rightarrow \Phi(\mathbf{n}), \quad \rho(\mathbf{x}) \rightarrow \rho(\mathbf{n}) \tag{30}
\end{equation*}
$$

## Second derivative discretised

Let's look at discretisation of derivatives again. Forward and backward derivatives:

$$
\begin{align*}
& \partial_{F} f(x)=\frac{f(x+a)-f(x)}{a}=f^{\prime}(x)+f^{\prime \prime}(x) \frac{a}{2}+\ldots=f^{\prime}(x)+O(a)  \tag{31}\\
& \partial_{B} f(x)=\frac{f(x)-f(x-a)}{a}=f^{\prime}(x)-f^{\prime \prime}(x) \frac{a}{2}+\ldots=f^{\prime}(x)+O(a)
\end{align*}
$$

Now we can combine the two:

$$
\begin{equation*}
\frac{\partial_{F}+\partial_{B}}{2} f(x)=f^{\prime}(x)+f^{\prime \prime \prime}(x) \frac{a^{2}}{3}+\ldots=f^{\prime}(x)+O\left(a^{2}\right) \tag{32}
\end{equation*}
$$

Choosing a symmetrised discretisation gives an extra order. (we also have $\left.\partial_{F} f(x)=f^{\prime}(x+a / 2)+O\left(a^{2}\right)\right)$
We can similarly show:

$$
\begin{align*}
\partial_{B} \partial_{F} f(x) & =\frac{f(x-a)-2 f(x)+f(x+a)}{a^{2}} \tag{33}
\end{align*}=f^{\prime \prime}(x)+O\left(a^{2}\right)
$$

## Discretised Poisson eq.

$$
\begin{equation*}
\sum_{i=1}^{d}\left(\Phi_{\mathbf{n}+\hat{i}}-2 \Phi_{\mathbf{n}}+\Phi_{\mathbf{n}-\hat{i}}\right)=-a^{2} \rho(\mathbf{n}) \tag{34}
\end{equation*}
$$

This a linear system of equations with $\rho$ giving the right hand side. matrix $A$ in one dimension is:

$$
\left(\begin{array}{ccccc}
-2 & 1 & 0 & \cdots & 0 \\
1 & -2 & 1 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 1 & -2 & 1 \\
0 & & \cdots & 1 & -2
\end{array}\right)
$$

In three dimensions:


We can also written this as:

$$
\begin{array}{r}
2 d \Phi_{\mathbf{n}}=\sum_{i=1}^{d}\left(\Phi_{\mathbf{n}+\hat{i}}+\Phi_{\mathbf{n}-\hat{i}}\right)+a^{2} \rho(\mathbf{n})  \tag{35}\\
\Phi_{\mathbf{n}}=\underbrace{\frac{1}{2 d} \sum_{i=1}^{d}\left(\Phi_{\mathbf{n}+\hat{i}}+\Phi_{\mathbf{n}-\hat{i}}\right)}_{\text {avr. of neighb. }}+\underbrace{\frac{a^{2}}{2 d} \rho(\mathbf{n})}_{\text {correction }}
\end{array}
$$

This suggests using a relaxation method: we start from some initial configuration and calculate averages until nothing changes

Jacobi method

$$
\begin{equation*}
\Phi_{\mathbf{n}}^{(t+1)}=\frac{1}{2 d} \sum_{i=1}^{d}\left(\Phi_{\mathbf{n}-\hat{i}}^{(t)}+\Phi_{\mathbf{n}+\hat{i}}^{(t)}\right)+\frac{a^{2}}{2 d} \rho(\mathbf{n}) \tag{36}
\end{equation*}
$$

Corresponds to the Jacobi iteration we had earlier

$$
\begin{equation*}
\mathbf{x}_{i+1}=D^{-1}\left(L \mathbf{x}_{i}+U \mathbf{x}_{i}+\mathbf{b}\right) \tag{37}
\end{equation*}
$$

We can use the values that we already have to make convergence faster:

$$
\begin{equation*}
\Phi_{\mathbf{n}}^{(t+1)}=\frac{1}{2 d} \sum_{i=1}^{d}\left(\Phi_{\mathbf{n}-\hat{i}}^{(t+1)}+\Phi_{\mathbf{n}+\hat{i}}^{(t)}\right)+\frac{a^{2}}{2 d} \rho(\mathbf{n}) \tag{38}
\end{equation*}
$$

Corresponds to the Gauss-Seidel iteration we had earlier (if the ordering of the elements $x_{i}$ of choosen right)

$$
\begin{equation*}
\mathbf{x}_{i+1}=D^{-1}\left(L \mathbf{x}_{i+1}+U \mathbf{x}_{i}+\mathbf{b}\right) \tag{39}
\end{equation*}
$$

No need to store $\phi^{(t+1)}$ separately. Also makes convergence faster.

Typically we need many steps in the same direction until the limiting value is reached. We could make convergence faster if we increase the change of the field artificially
$\rightarrow$ Successive overrelaxation

## Successive overrelaxation

$$
\begin{array}{r}
\Phi_{\mathbf{n}}^{\prime}=\frac{1}{2 d} \sum_{i=1}^{d}\left(\Phi_{\mathbf{n}-\hat{i}}^{(t+1)}+\Phi_{\mathbf{n}+\hat{i}}^{(t)}\right)+\frac{a^{2}}{2 d} \rho(\mathbf{n})  \tag{40}\\
\Phi_{\mathbf{n}}^{(t+1)}=\Phi_{\mathbf{n}}^{(t)}+\omega\left(\Phi_{\mathbf{n}}^{\prime}-\Phi_{\mathbf{n}}^{(t)}\right)=(1-\omega) \Phi_{\mathbf{n}}^{(t)}+\omega \Phi_{\mathbf{n}}^{\prime}
\end{array}
$$

Corresponds to the successive overrelaxation we had before:

$$
\begin{equation*}
\mathbf{x}_{i+1}=(1-\omega) x_{i}+\omega D^{-1}\left(L \mathbf{x}_{i+1}+U \mathbf{x}_{i}+\mathbf{b}\right) \tag{41}
\end{equation*}
$$

we again have:
$\omega=1$ : Gauss-Seidel
$\omega>1$ : overrelaxation
$\omega<1$ : underrelaxation
One can show that: 1 . the method is convergent for $0<\omega<2$, 2 . only $\omega>1$ can be faster than Gauss-Seidel (These are true for the Poisson problem, not generally for Gauss-Seidel with SOR)
The speed of convergence is quite sensitive to the value of $\omega$. experimentation is needed as the best $\omega$ value depends on the geometry of the problem


## Other boundary conditions

So far we use Dirichelet boundary conditions: $\Phi(x)$ on the boundary is given. We can also have: Neumann boundary conditions, where $\mathbf{n} \nabla \Phi(x)$ is given on the boundary, with $\mathbf{n}$ a normal vector of the boundary surface
We can also have: Robin boundary conditions, where $a \mathbf{n} \nabla \Phi(x)+b \Phi(x)$ is given on the boundary.

How do we implement Neumann numerically? In 1d:
We want to solve:

$$
\begin{equation*}
\triangle \Phi(x)=-\rho(x), \text { for } \Phi(i), \text { with } i=0, \ldots, N \tag{42}
\end{equation*}
$$

With the boundary condition $\partial_{x} \Phi(0)=d$ Let's introduce $\Phi(-1)$ "ghost point"
We than have:

$$
\begin{equation*}
\frac{\Phi(1)-\Phi(-1)}{2 a}=d, \quad \frac{\Phi(1)-2 \Phi(0)+\Phi(-1)}{a^{2}}=-\rho(0) \tag{43}
\end{equation*}
$$

$$
\begin{equation*}
\frac{\Phi(1)-\Phi(-1)}{2 a}=d, \quad \frac{\Phi(1)-2 \Phi(0)+\Phi(-1)}{a^{2}}=-\rho(0) \tag{44}
\end{equation*}
$$

We don't know what $\Phi(-1)$ is, but these two equations must be consistent:

$$
\begin{equation*}
\Phi(1)-2 a d=-a^{2} \rho(0)-\Phi(1)+2 \Phi(0) \tag{45}
\end{equation*}
$$

Which is equivalent to:

$$
\begin{equation*}
2 \Phi(1)-2 \Phi(0)=-a^{2} \rho(0)-2 a d \tag{46}
\end{equation*}
$$

So our matrix $A$ and right hand side is modified to:

$$
A=\left(\begin{array}{ccccc}
-2 & 2 & 0 & \ldots & 0 \\
1 & -2 & 1 & \ldots & 0 \\
0 & \ddots & \ddots & \ddots & \vdots \\
0 & \cdots & 1 & -2 & 1 \\
0 & & \cdots & 1 & -2
\end{array}\right), \quad b=\left(\begin{array}{c}
-a^{2} \rho(0)-2 a d \\
-a^{2} \rho(1) \\
\vdots \\
-a^{2} \rho(N-1) \\
-a^{2} \rho(N)
\end{array}\right)
$$

Robin boundary conditions can be dealt with similarly
Caveat: If there are Neumann boundary conditions everywhere in the system, $\Phi(x) \rightarrow \Phi(x)+$ const. is also a solution
$\Longrightarrow$ matrix $A$ will be singular. Simplest solution: fix $\Phi(0)=0$

## Finite Element Methods

What we have been doing so far is called "Finite difference method" There is a related method called Finite element method

We start from the reformulation of the given differential equation as a variational problem. This can be e.g. a minimization problem:

$$
\begin{equation*}
\text { Find } u \in V \text { such that } F(u) \leq F(v) \text { for all } v \in V \tag{47}
\end{equation*}
$$

where $V$ is a set of admissible functions and $F: V \rightarrow \mathbb{R}$ is a functional $v \in V$ can represent a continuously varying quantity such as temperature, electric potential, displacement in an elastic body, etc., $F(v)$ can be the total energy associated to $v$.

Usually $V$ is an infinite dimensional space. To be able to solve it, we restrict it to $V_{h} \subset V$, which consits of simple functions depending on finitely many parameters. The solution over $V_{h}$ consists of a large system of equations (linear on nonlinear).
E.g. $V_{h}$ can be picewise polynomial, this corresponds to the Ritz-Garlekin method.
for a FEM solution we need the following steps:
1 variational formulation of the PDE
2 discretisation using FEM: construction of the finite dimensional space $V_{h}$
3 deriving the equations describing the solution in $V_{h}$
4 Calculation (and codeing) using a computer

Sometimes there are multiple formulations depending on the choice of independent variables. The choice of $V_{h}$ is is influenced by the variational formulation, the accuracy requirement and the regularity properties of the solution
Advantages: Deals easily with complicated geometry, general boundary conditions and space dependent or non-linear material properties.

## Dielectric constant in Poisson eq.

Can we formulate the Poisson equation using a variational principle?

$$
\begin{equation*}
\triangle \Phi(x)=-\rho(x) \tag{48}
\end{equation*}
$$

Let's use:

$$
\begin{equation*}
F(\Phi)=\int d^{d} x\left(\frac{1}{2} \sum_{i}\left(\partial_{i} \Phi(x)\right)^{2}-\rho(x) \Phi(x)\right) \tag{49}
\end{equation*}
$$

We want to minimize this, we can write the Euler-Lagrange equations:

$$
\begin{equation*}
\sum_{i} \partial_{i} \frac{\delta F}{\delta \partial_{i} \Phi(x)}=\frac{\delta F}{\delta \Phi(x)} \Longrightarrow \triangle \Phi(x)=-\rho(x) \tag{50}
\end{equation*}
$$

We recognise $F(\Phi)$ as the energy of the electric field and the charge distribtuion in the potential.
What happens if we have a dielectric material:

$$
\begin{equation*}
\operatorname{div} D=\rho(x) \quad \Longrightarrow \quad \operatorname{div}(\epsilon(x) \nabla \phi)=-\rho(x) \tag{51}
\end{equation*}
$$

Easily incorporated in the variational principle:

$$
\begin{equation*}
F(\Phi)=\int d^{d} x\left(\frac{1}{2} \epsilon(x) \sum_{i}\left(\partial_{i} \Phi(x)\right)^{2}-\rho(x) \Phi(x)\right) \tag{52}
\end{equation*}
$$

Consider this in 1D:

$$
\begin{equation*}
F(\Phi)=\int d x\left(\frac{1}{2} \epsilon(x)\left(\partial_{x} \Phi(x)\right)^{2}-\rho(x) \Phi(x)\right) \tag{53}
\end{equation*}
$$

Now considering piecewise linear $\Phi$ :
(and slowly varying $\epsilon(x)$ and $\rho(x)$ )


$$
\begin{equation*}
F(\Phi)=\frac{1}{2 h} \sum_{k} \epsilon_{k+1 / 2}\left(\Phi_{k}-\Phi_{k+1}\right)^{2}-h \sum_{k} \rho_{k} \Phi_{k} \tag{54}
\end{equation*}
$$

$\partial F / \partial \Phi_{k}$ gives the minimum equations:

$$
\begin{equation*}
-\epsilon_{k+1 / 2} \Phi_{k+1}+\left(\epsilon_{k+1 / 2}+\epsilon_{k-1 / 2}\right) \Phi_{k}-\epsilon_{k-1 / 2} \Phi_{k-1}=h^{2} \rho_{k} \tag{55}
\end{equation*}
$$

Linear system of equations for $\Phi_{k}$

## Triangulation in FEM

Solving Poisson eq. in some 2d region:


The basis function $\varphi_{j}$.
Writing test functions $u_{h}=\sum_{i=k}^{N} \eta_{k} \varphi_{k}$ as piecewise linear on the triangles (therefore are fixed if we give the function on the vertices of the triangles)

$$
\begin{equation*}
F\left(u_{h}\right)=\int d^{d} x\left(\sum_{i}\left(\partial_{i} \sum_{k=1}^{N} \eta_{k} \varphi_{k}(x)\right)^{2}-\rho(x) \sum_{i=k}^{N} \eta_{k} \varphi_{k}(x)\right) \tag{56}
\end{equation*}
$$

The integrals can be calculated beforehand (amounts to integrating simple polynomials over triangle-shaped regions) and we get:

$$
\begin{equation*}
F\left(u_{h}\right)=\frac{1}{2} \eta_{i} A_{i j} \eta_{k}-\eta_{i} b_{i} \quad \Longrightarrow \quad A_{i k} \eta_{k}=b_{i} \tag{57}
\end{equation*}
$$

## Poisson eq: finite differences and FEM

Finite differences is a special case of FEM: using cubic lattice with piecewise linear functions is equivalent to finite differences


## Poisson eq. and Fourier Transformation

In continuum we can solve the poisson equation using Fourier equation:

$$
\begin{equation*}
\triangle \Phi(x)=-\rho(x) \underbrace{\Longrightarrow}_{F T}-k^{2} \Phi(k)=-\rho(k) \tag{58}
\end{equation*}
$$

Now we get $\Phi(x)$ as the inverse Fourier transformation of $\rho(k) / k^{2}$
We can do this on the lattice too:

$$
\triangle_{1 D}=\left(\begin{array}{ccccc}
-2 & 1 & 0 & \cdots & 1  \tag{59}\\
1 & -2 & 1 & \cdots & 0 \\
0 & \ddots & \ddots & \ddots & \vdots \\
\vdots & \ddots & \ddots & \ddots & \vdots \\
0 & & \ddots & \ddots & 1 \\
1 & 0 & \cdots & 1 & -2
\end{array}\right)=-k_{L A T}^{2}
$$

- We must use periodic boundary conditions in all dimensions
- total charge must be zero $(\rho(k=0)=0)$, otherwise we have a problem with Gauss law

