Numerical Methods
in Engineering and Applied Science

Lecture 4. Boundary-value problems for linear ODEs.

Let us now consider the following boundary-value problem:

$$\begin{aligned} &-\mu u''(x) + au'(x) + bu(x) = f(x), \quad x \in [0,1], \\ &u(0) = c, \quad u(1) = d \end{aligned} \tag{1}$$

where $\mu > 0$, $b \geq 0$. The boundary conditions are *Dirichlet* conditions. Let us introduce the following notation for the differential operator:

$$Lu(x) = -\mu u''(x) + au'(x) + bu(x). \tag{2}$$

The ODE becomes

$$Lu(x) = f(x). \tag{3}$$

Let us now introduce a central finite-difference approximation to the differential operator L,

$$L_h u(x) = -\mu \frac{u(x+h) - 2u(x) + u(x-h)}{h^2} + a\frac{u(x+h) - u(x-h)}{2h} + bu(x). \tag{4}$$

In general, $L_h u(x) \neq f(x)$, because $L_h u(x) \neq L(x)$ and, by definition, $Lu(x) = f(x)$.

The residual error

$$\tau(x) = L_h u(x) - Lu(x) \tag{5}$$

is called the **consistency error**. Let us analyze it.

The Taylor series expansion

$$
\begin{aligned}
u(x+h) &= u(x) + hu'(x) + \tfrac{h^2}{2}u''(x) + \tfrac{h^3}{6}u'''(x) + \tfrac{h^4}{24}u''''(x) + \mathcal{O}(h^5),\\
u(x-h) &= u(x) - hu'(x) + \tfrac{h^2}{2}u''(x) - \tfrac{h^3}{6}u'''(x) + \tfrac{h^4}{24}u''''(x) + \mathcal{O}(h^5)
\end{aligned} \tag{6}
$$

yields

$$
\begin{aligned}
\frac{u(x+h) - 2u(x) + u(x-h)}{h^2} &= u''(x) + \frac{h^2}{24}(u''''(\xi) + u''''(\zeta)),\\
\frac{u(x+h) - u(x-h)}{2h} &= u'(x) + \frac{h^2}{12}(u'''(\eta) + u'''(\delta)).
\end{aligned} \tag{7}
$$

where $\xi, \eta \in [x, x+h]$ and $\zeta, \delta \in [x-h, x]$.

Note that $|\frac{h^2}{24}(u''''(\xi) + u''''(\zeta))| <= \frac{h^2}{12}||u''''||_\infty$
and $|\frac{h^2}{12}(u'''(\eta) + u'''(\delta))| <= \frac{h^2}{6}||u'''||_\infty$
for any $x \in [0,1]$, where the L-infinity norm is defined as

$$
||u(x)||_\infty := \max_{x\in[0,1]} |u(x)|. \tag{8}
$$

We thus obtain an upper bound for the consistency error

$$||\tau||_\infty \leq \left(\frac{\mu}{12}||u''''||_\infty + \frac{a}{6}||u'''||_\infty\right) h^2. \tag{9}$$

The error decreases as h decreases, which means that the finite-difference approximation of the differential operator is **consistent**, i.e.,
$|L_h u(x) - Lu(x)| \to 0$ as $h \to 0$ for a function $u(x)$ at all points $x \in [0, 1]$.
The power exponent 2 in h^2 means that this is a second-order approximation.

Now let us use this consistency error bound to analyze the **convergence** of this finite-difference approximation. *We want to know if the finite-difference numerical solution of the boundary-value problem approaches the exact solution $u(x)$.*

Be reminded that the original BVP is

$$\begin{aligned} &-\mu u''(x) + au'(x) + bu(x) = f(x), \quad x \in [0,1], \\ &u(0) = c, \quad u(1) = d \end{aligned} \tag{10}$$

where $\mu > 0$, $a, b \geq 0$. To solve it numerically, we subdivide the domain with constant step $h = x_{i+1} - x_i$, $\forall i \in [0, N]$,

$$x_0 = 0 < x_1 < x_2 < ... < x_N < x_{N+1} = 1. \tag{11}$$

When we apply our finite-difference scheme to solve this BVP, we replace $Lu(x)$ in the l.h.s. of the equation by $L_h u_i$, and we introduce the notation $f_i = f(x_i)$. This gives a system of algebraic equations where u_i are the unknowns,

$$-\mu \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + a\frac{u_{i+1} - u_{i-1}}{2h} + bu_i = f_i, \quad i = 1, ..., N. \tag{12}$$

Let us rewrite it as

$$\begin{aligned} &u_{i-1}\left(-\frac{a}{2h} - \frac{\mu}{h^2}\right) + u_i\left(b + \frac{2\mu}{h^2}\right) + u_{i+1}\left(\frac{a}{2h} - \frac{\mu}{h^2}\right) = f_i, \quad i = 1, ..., N \\ &u_0 = c, \quad u_{N+1} = d \end{aligned}$$

The same is written in the matrix form as

$$A_h \boldsymbol{U} = \boldsymbol{f} \tag{13}$$

where

$$A_h = \begin{pmatrix} b + \frac{2\mu}{h^2} & \frac{a}{2h} - \frac{\mu}{h^2} & & & & & \\ -\frac{a}{2h} - \frac{\mu}{h^2} & b + \frac{2\mu}{h^2} & \frac{a}{2h} - \frac{\mu}{h^2} & & & & \\ & -\frac{a}{2h} - \frac{\mu}{h^2} & b + \frac{2\mu}{h^2} & \frac{a}{2h} - \frac{\mu}{h^2} & & & \\ & & & & \cdots & & \\ & & & & -\frac{a}{2h} - \frac{\mu}{h^2} & b + \frac{2\mu}{h^2} & \frac{a}{2h} - \frac{\mu}{h^2} \\ & & & & & -\frac{a}{2h} - \frac{\mu}{h^2} & b + \frac{2\mu}{h^2} \end{pmatrix}$$

$$\boldsymbol{U} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \\ \vdots \\ u_{N-1} \\ u_N \end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} f_1 + \left(\frac{a}{2h} + \frac{\mu}{h^2}\right) c \\ f_2 \\ f_3 \\ \vdots \\ f_{N-1} \\ f_N - \left(\frac{a}{2h} - \frac{\mu}{h^2}\right) d \end{pmatrix}$$

This tri-diagonal linear system can be solved by Gaussian elimination.

The same is written in the matrix form as

$$(-\mu D_h^{(2)} + aD_h + bI)\boldsymbol{U} = \boldsymbol{f} \tag{14}$$

where

$$D_h^{(2)} = \frac{1}{h^2}\begin{pmatrix} -2 & 1 & & & & & \\ 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & & & \cdots & & \\ & & & & 1 & -2 & 1 \\ & & & & & 1 & -2 \end{pmatrix}$$

$$D_h = \frac{1}{2h}\begin{pmatrix} 0 & 1 & & & & & \\ -1 & 0 & 1 & & & & \\ & -1 & 0 & 1 & & & \\ & & & & \cdots & & \\ & & & & -1 & 0 & 1 \\ & & & & & -1 & 0 \end{pmatrix}, \quad I = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & 1 & & & \\ & & & \cdots & & \\ & & & & 1 & \\ & & & & & 1 \end{pmatrix}$$

Note that matrices $D_h^{(2)}$ and D_h can be used to compute derivatives by matrix-vector multiplication when the values of $\boldsymbol{U}$ are given (with $u_0 = u_{N+1} = 0$).

Let us now examine the point-wise difference between the finite-difference numerical solution u_i and the exact solution $u(x_i)$ of the BVP,

$$e_i = u(x_i) - u_i, \tag{15}$$

with $e_0 = e_{N+1} = 0$ by definition. It is called the **discretization error**. If $||e_i|| \to 0$ as $h \to 0$, we say that the numerical solution **converges** to the exact solution. Let us check the convergence of our finite-difference solution. Using the definition of the consistency error, we write

$$\begin{aligned}
\tau_i &= L_h u(x_i) - f(x_i) = L_h u(x_i) - f_i = \\
&= \left(-\mu \frac{u(x_{i+1}) - 2u(x_i) + u(x_{i-1})}{h^2} + a\frac{u(x_{i+1}) - u(x_{i-1})}{2h} + bu(x_i)\right) \\
&- \left(-\mu \frac{u_{i+1} - 2u_i + u_{i-1}}{h^2} + a\frac{u_{i+1} - u_{i-1}}{2h} + bu_i\right) \\
&= -\mu \frac{e_{i+1} - 2e_i + e_{i-1}}{h^2} + a\frac{e_{i+1} - e_{i-1}}{2h} + be_i.
\end{aligned} \tag{16}$$

We then multiply the above result by e_i and sum up from 1 to N,

$$\sum_{i=0}^{N} \mu \frac{(e_{i+1} - e_i)^2}{h^2} + \sum_{i=1}^{N} b e_i^2 = \sum_{i=1}^{N} \tau_i e_i. \tag{17}$$

It follows that

$$\sum_{i=0}^{N} \mu \frac{(e_{i+1} - e_i)^2}{h^2} + \sum_{i=1}^{N} b e_i^2 \leq \sum_{i=1}^{N} |\tau_i||e_i| \tag{18}$$

and, consequently,

$$\sum_{i=0}^{N} \mu \frac{(e_{i+1} - e_i)^2}{h^2} \leq \sum_{i=1}^{N} |\tau_i||e_i|. \tag{19}$$

Further, we notice the identity

$$e_i = \sum_{j=0}^{i-1} (e_{j+1} - e_j) \tag{20}$$

from which it follows that

$$|e_i| \leq \sum_{j=0}^{N} |e_{j+1} - e_j|. \tag{21}$$

The Cauchy–Schwarz inequality gives us

$$\sum_{j=0}^{N} |e_{j+1} - e_j| \cdot 1 \leq \left(\sum_{j=0}^{N} (e_{j+1} - e_j)^2 \right)^{1/2} (N+1)^{1/2}. \tag{22}$$

Let m be the first occurrence of the maximum error,

$$m = \min \left\{ i \in 1, ..., N : |e_m| = \max_{j=1,...,N} |e_j| \right\}. \tag{23}$$

Combining the above inequalities and remembering that $h = 1/(N+1)$, we find

$$|e_m|^2 \le \left(\sum_{i=0}^{N}(e_{i+1}-e_i)^2\right)(N+1) \le \frac{h^2}{\mu}\sum_{i=1}^{N}|\tau_i|\ |e_i|\ (N+1) \le \frac{h}{\mu}|e_m|\sum_{i=1}^{N}|\tau_i|$$
$$\le \frac{h}{\mu}|e_m|N\left(\frac{\mu}{12}||u''''||_\infty + \frac{a}{6}||u'''||_\infty\right)h^2 \le |e_m|\left(\frac{1}{12}||u''''||_\infty + \frac{a}{6\mu}||u'''||_\infty\right)h^2,$$

It follows that

$$|e_m| \le \left(\frac{1}{12}||u''''||_\infty + \frac{a}{6\mu}||u'''||_\infty\right)h^2. \tag{24}$$

In other words, the L-infinity error norm is bounded as

$$||e||_\infty \le \left(\frac{1}{12}||u''''||_\infty + \frac{a}{6\mu}||u'''||_\infty\right)h^2. \tag{25}$$

We see that the *global discretization error* decreases as h^2, which means that our finite-difference solution is *second-order accurate*.

For the system $A_h\boldsymbol{U} = \boldsymbol{f}$ to have a solution, A_h should be non-singular.

One desirable property of the coefficients of matrix A_h is diagonal dominance. An $N \times N$ matrix $A = (a_{ij})$ is *strictly row diagonally dominant* if

$$|a_{ii}| > \sum_{j \neq i} |a_{ij}|, \quad i = 1, ..., N. \tag{26}$$

A is *strictly column diagonally dominant* if

$$|a_{ii}| > \sum_{j \neq i} |a_{ji}|, \quad i = 1, ..., N. \tag{27}$$

If $N \times N$ matrix A is strictly row or column diagonally dominant, then it is non-singular.

Proof: Assume that A is singular so that $A\mathbf{x} = 0$ for some non-zero $\mathbf{x}$. Let $|x_k| = \max\{|x_i| : i = 1, ..., N\}$. The k-th equation of $A\mathbf{x} = 0$ is

$$a_{kk}x_k = -\sum_{j\neq k} a_{kj}x_j. \tag{28}$$

It follows that

$$|a_{kk}||x_k| = |a_{kk}x_k| = \left|\sum_{j\neq k} a_{kj}x_j\right| \leq |x_k| \sum_{j\neq k} |a_{kj}|. \tag{29}$$

By assumption, $|x_k| > 0$ and we can divide the above by $|x_k|$ to obtain

$$|a_{kk}| \leq \sum_{j\neq k} |a_{kj}|. \tag{30}$$

This contradicts strict row diagonal dominance.
If A is strictly column diagonally dominant, then A^T is strictly row diagonally dominant, therefore it is non-singular. Then, A is also non-singular because $\det A = \det A^T \neq 0$.

For A_h to be strictly row (and column) diagonally dominant, we need

$$|b + \frac{2\mu}{h^2}| > |-\frac{a}{2h} - \frac{\mu}{h^2}| + |\frac{a}{2h} - \frac{\mu}{h^2}|. \tag{31}$$

This can be rewritten as

$$|2\mu + bh^2| > |\mu + \frac{ah}{2}| + |\mu - \frac{ah}{2}|. \tag{32}$$

Since we initially assumed that $\mu > 0$ and $b \geq 0$, the strict row diagonal dominance is achieved if h is sufficiently small. Specifically, if

$$h < 2\frac{\mu}{|a|}, \tag{33}$$

then strict diagonal dominance is achieved, because this guarantees

$$|2\mu + bh^2| > \mu + \frac{ah}{2} + \mu - \frac{ah}{2} = 2\mu. \tag{34}$$

If the diagonal dominance condition is not satisfied, A_h can be near singular. This means that small perturbation in $\boldsymbol{f}$ can lead to large errors in the solution $\boldsymbol{U}$. The diagonal dominance condition $h < 2\mu/|a|$ is rather stringent and can be avoided by using one-sided difference in place of central difference for the first derivative,

$$u'(x) = (u_{i+1} - u_i)/h \quad \text{if} \quad a < 0 \tag{35}$$

or

$$u'(x) = (u_i - u_{i-1})/h \quad \text{if} \quad a > 0. \tag{36}$$

These are often used in PDEs and they are called *upwind* differences in that context. The i-th row of the coefficient matrix A_h becomes

$$\begin{array}{llll} -\dfrac{\mu}{h^2} & -\dfrac{a}{h} + b + \dfrac{2\mu}{h^2} & \dfrac{a}{h} - \dfrac{\mu}{h^2} & \text{if} \quad a < 0, \\ -\dfrac{a}{h} - \dfrac{\mu}{h^2} & \dfrac{a}{h} + b + \dfrac{2\mu}{h^2} & -\dfrac{\mu}{h^2} & \text{if} \quad a > 0. \end{array} \tag{37}$$

Strict diagonal dominance holds independent of h. But this is a first-order accurate approximation.

Now let us consider the *Neumann* boundary conditions. Our BVP becomes

$$\begin{aligned} &-\mu u''(x) + a u'(x) + b u(x) = f(x), \quad x \in [0,1], \\ &u'(0) = c, \quad u'(1) = d \end{aligned} \tag{38}$$

In the difference equations, the values of u_0 and u_{N+1} are no longer known: they become unknowns. But we have $u'(0) = c$ and we can derive additional equations. Let us approximate u'' at $x = 0$ by

$$u''(0) = \frac{1}{h^2}(u_{-1} - 2u_0 + u_1) \tag{39}$$

using a grid point $-h$ outside the interval. Then by the boundary condition

$$c = u'(0) = \frac{1}{2h}(u_1 - u_{-1}) \tag{40}$$

we have $u_{-1} = u_1 - 2ch$, and we can use this to eliminate u_{-1}:

$$u''(0) = \frac{1}{h^2}(2u_1 - 2u_0 - 2ch) \tag{41}$$

This approximation is first-order accurate in general, and second-order accurate if $c = 0$.

At $x = 1$ we obtain

$$u''(1) = \frac{1}{h^2}(2u_N - 2u_{N+1} + 2dh. \tag{42}$$

We thus obtain two finite-difference equations to be included in the system:

$$\left(b + \frac{2\mu}{h^2}\right) u_0 - \frac{2\mu}{h^2} u_1 = f_0 - ac + \frac{2c\mu}{h} \tag{43}$$

and

$$\left(b + \frac{2\mu}{h^2}\right) u_{N+1} - \frac{2\mu}{h^2} u_N = f_{N+1} - ad + \frac{2d\mu}{h}. \tag{44}$$

We now have a linear system of $N + 2$ equations to solve. A_h becomes an $(N + 2) \times (N + 2)$ matrix.

$$A_h \boldsymbol{U} = \boldsymbol{f}, \quad \text{where} \tag{45}$$

$$A_h = \begin{pmatrix} b+\frac{2\mu}{h^2} & -\frac{2\mu}{h^2} & & & & & \\ -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} & & & & \\ & -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} & & & \\ & & & & \cdots & & \\ & & & & -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} \\ & & & & & -\frac{2\mu}{h^2} & b+\frac{2\mu}{h^2} \end{pmatrix}$$

$$\boldsymbol{U} = \begin{pmatrix} u_0 \\ u_1 \\ u_2 \\ \vdots \\ u_N \\ u_{N+1} \end{pmatrix}, \quad \boldsymbol{f} = \begin{pmatrix} f_0 - ac + \frac{2c\mu}{h} \\ f_1 \\ f_2 \\ \vdots \\ f_N \\ f_{N+1} - ad + \frac{2d\mu}{h} \end{pmatrix}$$

Note that it is also possible to have a Dirichlet boundary condition on one end of the interval and Neumann condition on the other end. In that situation, the first and last equations from (13) and (45) are used accordingly.
Finite-difference approximation for the Robin boundary conditions

$$\eta_1 u(0) + \eta_2 u'(0) = c, \qquad \gamma_1 u(1) + \gamma_2 u'(1) = d \tag{46}$$

is derived similarly.
If $a = b = 0$, the matrix A_h of (45) becomes

$$A_h = \frac{1}{h^2} \begin{pmatrix} 2 & -2 & & & & \\ -1 & 2 & -1 & & & \\ & & & \cdots & & \\ & & & & -2 & 2 \end{pmatrix}$$

We see that $A_h \boldsymbol{e} = \mathbf{0}$ for $\boldsymbol{e} = (1, 1, ..., 1)^T$. Thus, this A_h *is singular!* This reflects the non-uniqueness of the underlying exact solution of (38). The solution is also not unique if $a \neq 0$ but $b = 0$.

Similar non-uniqueness situation happens when using *periodic* boundary conditions, such that $u_0 = u_N$ and $u_{N+1} = u_1$. In that case we obtain a system of N equations with the matrix

$$A_h = \begin{pmatrix} b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} & & & & & -\frac{a}{2h}-\frac{\mu}{h^2} \\ -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} & & & & \\ & -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} & & & \\ & & & \dots & & & \\ & & & & -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} & \frac{a}{2h}-\frac{\mu}{h^2} \\ \frac{a}{2h}-\frac{\mu}{h^2} & & & & & -\frac{a}{2h}-\frac{\mu}{h^2} & b+\frac{2\mu}{h^2} \end{pmatrix} \tag{47}$$

This matrix is singular if $b = 0$.
A unique solution can be obtained if we replace one of the equations by

$$u_1 + u_2 + ... + u_N = const \tag{48}$$

The value of *const* will control the mean value of the solution.