Numerical Methods for PDEs

Local Truncation Error, Consistency, and Matrix Version of the FTCS Scheme

(Lecture 4, Week 2)

Markus Schmuck

Department of Mathematics and Maxwell Institute for Mathematical Sciences Heriot-Watt University, Edinburgh

Edinburgh, 19 January, 2015







3 Matrix formulation of the FTCS scheme





M. Schmuck (Heriot-Watt University) Numerical Methods for PDEs, Lecture 4

First examination: The Local Truncation Error (LTE) Preparatory work:

1. Rewrite the PDE (heat equation) in operator form

$$Lu = 0$$
, with $L = \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}$

2. FD approximation of L: For the FTCS scheme we have

 $L_{k,h}=D_t^+-D_x^2.$

Recall that the numerical solution solves

$$0 = L_{k,h} w_j^n = \frac{w_j^{n+1} - w_j^n}{k} - \frac{w_{j-1}^n - 2w_j^n + w_{j+1}^n}{h^2}.$$



First examination: The Local Truncation Error (LTE) Preparatory work:

1. Rewrite the PDE (heat equation) in operator form

$$Lu = 0$$
, with $L = \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}$

2. FD approximation of L: For the FTCS scheme we have

 $L_{k,h}=D_t^+-D_x^2.$

Recall that the numerical solution solves

$$0 = L_{k,h} w_j^n = \frac{w_j^{n+1} - w_j^n}{k} - \frac{w_{j-1}^n - 2w_j^n + w_{j+1}^n}{h^2}$$



First examination: The Local Truncation Error (LTE) Preparatory work:

1. Rewrite the PDE (heat equation) in operator form

$$Lu = 0$$
, with $L = \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}$

2. FD approximation of L: For the FTCS scheme we have

$$L_{k,h}=D_t^+-D_x^2.$$

Recall that the numerical solution solves

$$0 = L_{k,h} w_j^n = \frac{w_j^{n+1} - w_j^n}{k} - \frac{w_{j-1}^n - 2w_j^n + w_{j+1}^n}{h^2}$$



First examination: The Local Truncation Error (LTE) Preparatory work:

1. Rewrite the PDE (heat equation) in operator form

$$Lu = 0$$
, with $L = \frac{\partial}{\partial t} - \frac{\partial^2}{\partial x^2}$

2. FD approximation of L: For the FTCS scheme we have

$$L_{k,h}=D_t^+-D_x^2.$$

Recall that the numerical solution solves

$$0 = L_{k,h} w_j^n = \frac{w_j^{n+1} - w_j^n}{k} - \frac{w_{j-1}^n - 2w_j^n + w_{j+1}^n}{h^2}.$$

Definition. Let $L_{k,h}$ be the difference operator approximating *L*. The local truncation error (*LTE*) is given by the leading terms of the Taylor expansion of $L_{k,h}u(x_j, t_n) = 0$, where *u* satisfies Lu = 0.

In other words: Apply the finite difference operator $L_{k,h}$ of the PDE on the exact solution u, then Taylor expand and cancel as many terms as possible by using Lu = 0 for instance. That means $LTE = LOT [L_{k,h}u] (x_j, t_n)$, where $LOT [\cdot]$ denotes leading order Taylor expansion.



Definition. Let $L_{k,h}$ be the difference operator approximating *L*. The local truncation error (*LTE*) is given by the leading terms of the Taylor expansion of $L_{k,h}u(x_i, t_n) = 0$, where *u* satisfies Lu = 0.

In other words: Apply the finite difference operator $L_{k,h}$ of the PDE on the exact solution u, then Taylor expand and cancel as many terms as possible by using Lu = 0 for instance. That means $LTE = LOT [L_{k,h}u] (x_j, t_n)$, where $LOT [\cdot]$ denotes leading order Taylor expansion.



Example: LTE of the FTCS scheme

Calculate the LTE of the FTCS scheme of the heat equation:

$$LTE = L_{k,h}u(x_j, t_n)$$

= $\frac{u(x_j, t_n + k) - u(x_j, t_n)}{k}$
 $- \frac{u(x_j - h, t_n) - 2u(x_j, t_n) + u(x_j + h, t_n)}{h^2}$



Example: LTE of the FTCS scheme

Calculate the LTE of the FTCS scheme of the heat equation:

$$\begin{aligned} \text{TTE} &= L_{k,h} u(x_j, t_n) \\ &= \frac{u(x_j, t_n + k) - u(x_j, t_n)}{k} \\ &- \frac{u(x_j - h, t_n) - 2u(x_j, t_n) + u(x_j + h, t_n)}{h^2} \\ &= \frac{1}{k} \left(u + ku_t + \frac{1}{2}k^2 u_{tt} + \dots - u \right) \Big|_{x_j, t_n} \\ &- \frac{1}{h^2} \left(u - hu_x + \frac{1}{2}h^2 u_{xx} - \frac{1}{3!}h^3 u_{xxx} + \frac{1}{4!}h^4 u_{xxxx} - 2u + u + hu_x + \frac{1}{2}h^2 u_{xx} + \frac{1}{3!}h^3 u_{xxx} + \frac{1}{4!}h^4 u_{xxxx} \right) \Big|_{x_i, t_n} \end{aligned}$$



Example: LTE of the FTCS scheme

Calculate the LTE of the FTCS scheme of the heat equation:

$$\begin{aligned} \text{LTE} &= L_{k,h} u(x_j, t_n) \\ &= \frac{u(x_j, t_n + k) - u(x_j, t_n)}{k} \\ &- \frac{u(x_j - h, t_n) - 2u(x_j, t_n) + u(x_j + h, t_n)}{h^2} \\ &= \frac{1}{k} \left(u + ku_t + \frac{1}{2}k^2 u_{tt} + \dots - u \right) \Big|_{x_j, t_n} \\ &- \frac{1}{h^2} \left(u - hu_x + \frac{1}{2}h^2 u_{xx} - \frac{1}{3!}h^3 u_{xxx} + \frac{1}{4!}h^4 u_{xxxx} - 2u + u + hu_x + \frac{1}{2}h^2 u_{xx} + \frac{1}{3!}h^3 u_{xxx} + \frac{1}{4!}h^4 u_{xxxx} \right) \Big|_{x_j, t_n} \\ &= u_t(x_j, t_n) + \frac{1}{2}ku_{tt}(x_j, t_n) - u_{xx}(x_j, t_n) - \frac{1}{12}h^2 u_{xxxx}(x_j, t_n) + \dots \end{aligned}$$

LTE =
$$\frac{k}{2}u_{tt} - \frac{h^2}{12}u_{xxxx} + \mathcal{O}(k^2, h^4)$$
.

By using again the assumption on smoothness and the heat equation, that is,

$$u_{tt} = \frac{\partial}{\partial t} u_t = \frac{\partial}{\partial t} u_{xx} = \frac{\partial^2}{\partial x^2} u_t = \frac{\partial^2}{\partial x^2} u_{xx} = u_{xxxx} \,,$$

we can rewrite the LTE as

$$\text{LTE} = \left(\frac{k}{2} - \frac{h^2}{12}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) = \frac{h^2}{2} \left(r - \frac{1}{6}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) \,,$$

where $r = k/h^2$. Hence, if r = 1/6, then the $\mathcal{O}(h^2)$ terms vanish and the LTE becomes $\mathcal{O}(k^2, h^4)$. The *leading term* of the LTE (in green above) is *first order accurate in time* and *2nd order accurate in space*

$$\text{LTE} = \frac{k}{2}u_{tt} - \frac{h^2}{12}u_{xxxx} + \mathcal{O}(k^2, h^4).$$

By using again the assumption on smoothness and the heat equation, that is,

$$u_{tt} = \frac{\partial}{\partial t} u_t = \frac{\partial}{\partial t} u_{xx} = \frac{\partial^2}{\partial x^2} u_t = \frac{\partial^2}{\partial x^2} u_{xx} = u_{xxxx},$$

we can rewrite the LTE as

$$\text{LTE} = \left(\frac{k}{2} - \frac{h^2}{12}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) = \frac{h^2}{2} \left(r - \frac{1}{6}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) \,,$$

where $r = k/h^2$. Hence, if r = 1/6, then the $\mathcal{O}(h^2)$ terms vanish and the LTE becomes $\mathcal{O}(k^2, h^4)$. The *leading term* of the LTE (in green above) is *first order accurate in time* and *2nd order accurate in space*

$$\text{LTE} = \frac{k}{2}u_{tt} - \frac{h^2}{12}u_{xxxx} + \mathcal{O}(k^2, h^4).$$

By using again the assumption on smoothness and the heat equation, that is,

$$u_{tt} = \frac{\partial}{\partial t} u_t = \frac{\partial}{\partial t} u_{xx} = \frac{\partial^2}{\partial x^2} u_t = \frac{\partial^2}{\partial x^2} u_{xx} = u_{xxxx} ,$$

we can rewrite the LTE as

$$LTE = \left(\frac{k}{2} - \frac{h^2}{12}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) = \frac{h^2}{2} \left(r - \frac{1}{6}\right) u_{xxxx} + \mathcal{O}(k^2, h^4),$$

where $r = k/h^2$. Hence, if r = 1/6, then the $O(h^2)$ terms vanish and the LTE becomes $O(k^2, h^4)$. The *leading term* of the LTE (in green above) is *first order accurate in time* and *2nd order accurate in space*.

$$\text{LTE} = \frac{k}{2}u_{tt} - \frac{h^2}{12}u_{xxxx} + \mathcal{O}(k^2, h^4).$$

By using again the assumption on smoothness and the heat equation, that is,

$$u_{tt} = \frac{\partial}{\partial t} u_t = \frac{\partial}{\partial t} u_{xx} = \frac{\partial^2}{\partial x^2} u_t = \frac{\partial^2}{\partial x^2} u_{xx} = u_{xxxx} ,$$

we can rewrite the LTE as

$$LTE = \left(\frac{k}{2} - \frac{h^2}{12}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) = \frac{h^2}{2} \left(r - \frac{1}{6}\right) u_{xxxx} + \mathcal{O}(k^2, h^4),$$

where $r = k/h^2$. Hence, if r = 1/6, then the $\mathcal{O}(h^2)$ terms vanish and the LTE becomes $\mathcal{O}(k^2, h^4)$. The leading term of the LTE (in green above) is first order accurate in time and 2nd order accurate in space.

$$\text{LTE} = \frac{k}{2}u_{tt} - \frac{h^2}{12}u_{xxxx} + \mathcal{O}(k^2, h^4).$$

By using again the assumption on smoothness and the heat equation, that is,

$$u_{tt} = \frac{\partial}{\partial t} u_t = \frac{\partial}{\partial t} u_{xx} = \frac{\partial^2}{\partial x^2} u_t = \frac{\partial^2}{\partial x^2} u_{xx} = u_{xxxx} ,$$

we can rewrite the LTE as

$$LTE = \left(\frac{k}{2} - \frac{h^2}{12}\right) u_{xxxx} + \mathcal{O}(k^2, h^4) = \frac{h^2}{2} \left(r - \frac{1}{6}\right) u_{xxxx} + \mathcal{O}(k^2, h^4),$$

where $r = k/h^2$. Hence, if r = 1/6, then the $\mathcal{O}(h^2)$ terms vanish and the LTE becomes $\mathcal{O}(k^2, h^4)$. The *leading term* of the LTE (in green above) is *first order accurate in time* and *2nd order accurate in space*.



Definition. Let LTE be the local truncation error of a scheme. A scheme is said to be consistent, if LTE \rightarrow 0 as $h, k \rightarrow 0$. Moreover, a scheme is said to be of order p in space and q in time, i the LTE is of order $O(h^p, k^q)$.

Remark. If *r* is a fixed value, then we can rewrite the LTE as $O(h^p)$ as shown above, i.e., the FTCS scheme is 2nd order if $r \neq 1/6$ and 4th order if r = 1/6.



Definition. Let LTE be the local truncation error of a scheme. A scheme is said to be consistent, if LTE $\rightarrow 0$ as $h, k \rightarrow 0$. Moreover, a scheme is said to be of order p in space and q in time, if the LTE is of order $\mathcal{O}(h^p, k^q)$.

Remark. If *r* is a fixed value, then we can rewrite the LTE as $O(h^p)$ as shown above, i.e., the FTCS scheme is 2nd order if $r \neq 1/6$ and 4th order if r = 1/6.



Definition. Let LTE be the local truncation error of a scheme. A scheme is said to be consistent, if LTE $\rightarrow 0$ as $h, k \rightarrow 0$. Moreover, a scheme is said to be of order p in space and q in time, if the LTE is of order $O(h^p, k^q)$.

Remark. If *r* is a fixed value, then we can rewrite the LTE as $O(h^p)$ as shown above, i.e., the FTCS scheme is 2nd order if $r \neq 1/6$ and 4th order if r = 1/6.



FTCS scheme in matrix form

Basic definitions:

- Vector containing numerical solutions w_i^n at internal grid points
- j = 1, 2, ..., J 1 at time t_n

$$\mathbf{w}^n = \begin{pmatrix} w_1^n \\ w_2^n \\ \vdots \\ w_{J-1}^n \end{pmatrix}$$

• Vector of initial values (known)

$$\mathbf{w}^{0} = \mathbf{u}^{0} = \begin{pmatrix} F(x_{1}) \\ F(x_{2}) \\ \vdots \\ F_{J-1} \end{pmatrix}$$



FTCS scheme in matrix form

Basic definitions:

- Vector containing numerical solutions w_i^n at internal grid points
- j = 1, 2, ..., J 1 at time t_n

$$\mathbf{w}^n = \begin{pmatrix} w_1^n \\ w_2^n \\ \vdots \\ w_{J-1}^n \end{pmatrix}$$

• Vector of initial values (known)

$$\mathbf{w}^{0} = \mathbf{u}^{0} = \begin{pmatrix} F(x_{1}) \\ F(x_{2}) \\ \vdots \\ F_{J-1} \end{pmatrix}$$

These definitions allow us to rewrite the FTCS scheme

$$w_j^{n+1} = rw_{j-1}^n + (1-2r)w_j^n + rw_{j+1}^n, \quad j = 1, \ldots, J-1,$$

as follows

 $\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n$,

where





These definitions allow us to rewrite the FTCS scheme

$$w_j^{n+1} = rw_{j-1}^n + (1-2r)w_j^n + rw_{j+1}^n, \quad j = 1, \ldots, J-1,$$

as follows

 $\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n \,,$

where





These definitions allow us to rewrite the FTCS scheme

$$w_j^{n+1} = rw_{j-1}^n + (1-2r)w_j^n + rw_{j+1}^n, \quad j = 1, \ldots, J-1,$$

as follows

$$\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n$$
,

where





$$\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n$$

= $S(S\mathbf{w}^{n-1} + \mathbf{b}^{n-1}) + \mathbf{b}^n$
= $S^2\mathbf{w}^{n-1} + S\mathbf{b}^{n-1} + \mathbf{b}^n$,

gives by repeating back to \mathbf{w}^0 (or by induction)

 $\mathbf{w}^{n+1} = S^{n+1}\mathbf{w}^0 + S^n\mathbf{b}^0 + S^{n-1}\mathbf{b}^1 + \dots + S\mathbf{b}^{n-1} + \mathbf{b}^n,$ here $S^n := \underbrace{S \times S \times S \times \dots \times S}_{n \text{ times}}.$

Remark. If we have homogeneous Dirichlet BC, i.e., $\alpha(t) = \beta(t) = 0$, then **b**^{*n*} = **0** for all *n* and hence

$$\mathbf{w}^{n+1} = S^{n+1} \mathbf{w}^0$$

$$\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n$$

= $S(S\mathbf{w}^{n-1} + \mathbf{b}^{n-1}) + \mathbf{b}^n$
= $S^2\mathbf{w}^{n-1} + S\mathbf{b}^{n-1} + \mathbf{b}^n$,

gives by repeating back to w⁰ (or by induction)

 $\mathbf{w}^{n+1} = S^{n+1}\mathbf{w}^0 + S^n\mathbf{b}^0 + S^{n-1}\mathbf{b}^1 + \dots + S\mathbf{b}^{n-1} + \mathbf{b}^n,$ where $S^n := \underbrace{S \times S \times S \times \dots \times S}_{n \text{ times}}.$

Remark. If we have homogeneous Dirichlet BC, i.e., $\alpha(t) = \beta(t) = 0$, then **b**^{*n*} = **0** for all *n* and hence

 $\mathbf{w}^{n+1} = S^{n+1} \mathbf{w}^0$



$$\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n$$

= $S(S\mathbf{w}^{n-1} + \mathbf{b}^{n-1}) + \mathbf{b}^n$
= $S^2\mathbf{w}^{n-1} + S\mathbf{b}^{n-1} + \mathbf{b}^n$,

gives by repeating back to w⁰ (or by induction)

 $\mathbf{w}^{n+1} = S^{n+1}\mathbf{w}^0 + S^n\mathbf{b}^0 + S^{n-1}\mathbf{b}^1 + \dots + S\mathbf{b}^{n-1} + \mathbf{b}^n,$ where $S^n := \underbrace{S \times S \times S \times \dots \times S}_{n \text{ times}}.$

Remark. If we have homogeneous Dirichlet BC, i.e., $\alpha(t) = \beta(t) = 0$, then **b**^{*n*} = **0** for all *n* and hence

 $\mathbf{w}^{n+1} = S^{n+1} \mathbf{w}^0$



$$\mathbf{w}^{n+1} = S\mathbf{w}^n + \mathbf{b}^n$$

= $S(S\mathbf{w}^{n-1} + \mathbf{b}^{n-1}) + \mathbf{b}^n$
= $S^2\mathbf{w}^{n-1} + S\mathbf{b}^{n-1} + \mathbf{b}^n$,

gives by repeating back to w^0 (or by induction)

 $\mathbf{w}^{n+1} = S^{n+1}\mathbf{w}^0 + S^n\mathbf{b}^0 + S^{n-1}\mathbf{b}^1 + \dots + S\mathbf{b}^{n-1} + \mathbf{b}^n,$ where $S^n := \underbrace{S \times S \times S \times \dots \times S}_{n \text{ times}}.$

Remark. If we have homogeneous Dirichlet BC, i.e., $\alpha(t) = \beta(t) = 0$, then $\mathbf{b}^n = \mathbf{0}$ for all *n* and hence

$$\mathbf{w}^{n+1} = S^{n+1} \mathbf{w}^0 \, .$$



Notation in matlab

Mathematical (analytical) notation:

The spatial index *j* goes over the spatial grid points 0, 2, ..., J (including boundary points).

Notation adapted to matlab:

The spatial index j goes over the spatial grid points 1, 2, ..., J + 1 (including boundary points).

Sparse matrix:

Declare the matrix S as sparse in matlab (as most elements are zero)

$$\begin{split} S &= sparse(diag((1-2*r)*ones(J-1,1)) \dots \\ &+ diag(r*ones(J-2,1),1) + diag(r*ones(J-2,1),-1)); \end{split}$$

"sparse" squeezes out any zero elements (type "help sparse" in matlab shell for more information)





Notation in matlab

Mathematical (analytical) notation:

The spatial index *j* goes over the spatial grid points $0, 2, \ldots, J$ (including boundary points).

Notation adapted to matlab:

The spatial index *j* goes over the spatial grid points 1, 2, ..., J + 1 (including boundary points).

Sparse matrix:

Declare the matrix S as sparse in matlab (as most elements are zero)

$$\begin{split} S &= sparse(diag((1-2*r)*ones(J-1,1)) \dots \\ &+ diag(r*ones(J-2,1),1) + diag(r*ones(J-2,1),-1)); \end{split}$$

"sparse" squeezes out any zero elements (type "help sparse" in matlab shell for more information)





Notation in matlab

Mathematical (analytical) notation:

The spatial index *j* goes over the spatial grid points 0, 2, ..., J (including boundary points).

Notation adapted to matlab:

The spatial index *j* goes over the spatial grid points 1, 2, ..., J + 1 (including boundary points).

Sparse matrix:

Declare the matrix S as sparse in matlab (as most elements are zero)

$$\begin{split} S &= \text{sparse}(\ \text{diag}((1-2*r)*\text{ones}(J-1,1)) \ \dots \\ &+ \text{diag}(r*\text{ones}(J-2,1),1) + \text{diag}(r*\text{ones}(J-2,1),-1) \); \end{split}$$

"sparse" squeezes out any zero elements (type "help sparse" in matlab shell for more information)





Summary of learning targets:

1. What means LTE and how can it be obtained for a finite difference scheme?

2. How does the finite difference operator for the FTCS scheme with respect to all internal grid points look like? Does it show a special structure?

3. Can you define diagonal matrices in matlab? How can one prevent the storage of zeros?



1. What means LTE and how can it be obtained for a finite difference scheme?

2. How does the finite difference operator for the FTCS scheme with respect to all internal grid points look like? Does it show a special structure?

3. Can you define diagonal matrices in matlab? How can one prevent the storage of zeros?



1. What means LTE and how can it be obtained for a finite difference scheme?

2. How does the finite difference operator for the FTCS scheme with respect to all internal grid points look like? Does it show a special structure?

3. Can you define diagonal matrices in matlab? How can one prevent the storage of zeros?

