Recap:
Two-point BVP, finite difference methods for BVP;
Fully implicit RK methods, collocation methods for BVP;
Richardson extrapolation, Romberg integration

Numerical solution of the heat equation
IBVP (initial boundary value problem) of the heat equation:
\[
\begin{cases}
  u_t = u_{xx} \\
  u(x, t_0) = f(x) \\
  u(a, t) = g_1(x), \quad u(b, t) = g_2(x)
\end{cases}
\]

Numerical grid:
\[\Delta x = \frac{b-a}{N+1}, \quad x_i = a + i \Delta x\]
\[x_0 = a, \quad x_{N+1} = b, \quad \{x_i, 1 \leq i \leq N\} \text{ are internal points.}\]
\[t_n = t_0 + n \Delta t\]

Notation:
\[u(x_i, t_n^1): \text{ exact solution at } (x_i, t_n)\]
\[u_i^n: \text{ numerical approximation of } u(x_i, t_n)\]

Discretization of derivatives:
\[
\left. u_{xx} \right|_{(x_i, t_n)} \approx \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}
\]
\[
\left. u_t \right|_{(x_i, t_n)} \approx \frac{u_i^{n+1} - u_i^n}{\Delta t}
\]

Discretization of PDE:
\[
\frac{u_i^{n+1} - u_i^n}{\Delta t} = \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{(\Delta x)^2}
\]

This is called the FTCS method (Forward-time, Central-space method).
Let \( r = \frac{\Delta t}{(\Delta x)^2} \). We write out the FTSC method along with the initial and boundary conditions, which are needed for closing the system over finite domain.

The FTCS method for IBVP:

\[
\begin{align*}
    u_{i}^{n+1} &= u_{i}^{n} + r \left( u_{i+1}^{n} - 2u_{i}^{n} + u_{i-1}^{n} \right), \quad 1 \leq i \leq N, \quad r = \frac{\Delta t}{(\Delta x)^2} \\
    u_{0}^{n} &= f(x_{i}), \quad 1 \leq i \leq N \\
    u_{0}^{n} &= g_{1}(t_{n}), \quad u_{N+1}^{n} = g_{2}(t_{n}), \quad n \geq 0
\end{align*}
\]

Sometimes, for theoretical simplicity and convenience, we also consider initial value problems over the infinite domain. But keep in mind that in computer implementation and in real applications, only IBVPs make practical sense.

IVP (initial value problem) of the heat equation:

\[
\begin{align*}
    u_{t} &= u_{xx} \\
    u(x, t_{0}) &= f(x)
\end{align*}
\]

The FTCS method for IVP:

\[
\begin{align*}
    u_{i}^{n+1} &= u_{i}^{n} + r \left( u_{i+1}^{n} - 2u_{i}^{n} + u_{i-1}^{n} \right), \quad -\infty < i < +\infty, \quad r = \frac{\Delta t}{(\Delta x)^2} \\
    u_{i}^{0} &= f(x_{i}), \quad -\infty < i < +\infty
\end{align*}
\]

Next, we introduce

Local truncation error
Consistency, order of accuracy
Stability
Global error
Convergence

Local truncation error:

When we substitute an exact solution into the numerical method, the residual term is called the local truncation error (LTE) and is denoted by \( e_{i}^{n} (\Delta x, \Delta t) \).

Consistency:
Suppose a numerical method satisfies
The FTCS method is consistent with the heat equation.

The FTCS method:

\[
u_{i+1}^n - u_i^n - r(u_{i+1}^n - 2u_i^n + u_{i-1}^n) = 0, \quad r = \frac{\Delta t}{(\Delta x)^2}
\]

Its local truncation error is defined as

\[
e_i^n(\Delta x, \Delta t) = u(x_i, t_{n+1}) - u(x_i, t_n) - r(u(x_{i+1}, t_n) - 2u(x_i, t_n) + u(x_{i-1}, t_n))
\]

Expanding everything around \((x_i, t_n)\) gives us

\[
u(x_i, t_{n+1}) - u(x_i, t_n) = u_i \Delta t + \frac{1}{2} u_{ii} \Delta t^2 + o(\Delta t^2)
\]

\[
u(x_{i+1}, t_n) - 2u(x_i, t_n) + u(x_{i-1}, t_n) = 2 \cdot \frac{1}{2!} u_{xx} (\Delta x)^2 + 2 \cdot \frac{1}{4!} u_{xxxx} (\Delta x)^4 + o(\Delta x)^4)
\]

Using \(u_t = u_{xx}, \ u_{tt} = u_{xxxx}\) and \(r = \frac{\Delta t}{(\Delta x)^2}\), we write the local truncation error as:

\[
e_i^n(\Delta x, \Delta t) = \Delta t \left[ u_{xx} + \frac{1}{2} u_{xxxx} \Delta t + o(\Delta t^2) \right] - \Delta t \left[ u_{xx} + \frac{1}{12} u_{xxxx} (\Delta x)^2 + o(\Delta x)^2 \right]
\]

\[
e_i^n(\Delta x, \Delta t) = \Delta t u_{xxxx} \left[ \frac{1}{2} \Delta t - \frac{1}{12} (\Delta x)^2 + \cdots \right]
\]

\[
e_i^n(\Delta x, \Delta t) \Delta t = u_{xxxx} \left[ \frac{1}{2} \Delta t - \frac{1}{12} (\Delta x)^2 + \cdots \right] \rightarrow 0 \quad \text{as} \quad (\Delta x, \Delta t) \rightarrow 0
\]

The FTCS is consistent. It is first order in time and second order in space.
The BTCS method (Backward-time, Central-space) for IVP

\[ u_i^{n+1} = u_i^n + r \left( u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1} \right), \quad -\infty < i < +\infty, \quad r = \frac{\Delta t}{(\Delta x)^2} \]

\[ u_i^0 = f(x_i), \quad -\infty < i < +\infty \]

Its local truncation error has the expression:

\[ e_i^n(\Delta x, \Delta t) = \Delta t u_xxxx \left[ -\frac{1}{2} \Delta t - \frac{1}{12} (\Delta x)^2 + \cdots \right] \]

where all expansions are carried out around \((x_i, t_{n+1})\).

The BTCS is consistent. It is first order in time and second order in space.

Linear operator notation:

We write a general linear numerical method as a linear operator.

\[ u^{n+1} = L_{\text{num}}(u^n) \]

where

\[ L_{\text{num}} = \text{linear operator representing the method} \]

\[ u^n = \{u_i^n, \ 1 \leq i \leq N\} \text{ in an IBVP} \]

Or \[ u^n = \{u_i^n, \ -\infty < i < +\infty\} \text{ in an IVP} \]

With this short notation, the local truncation error is

\[ \left\{ e_i^n(\Delta x, \Delta t) \right\} = \left\{ u(x_i, t_{n+1}) \right\} - L_{\text{num}} \left\{ u(x_i', t_n) \right\} \]

This is also the error of one step.

After advancing \(n\) steps in time, the numerical solution is

\[ u^n = (L_{\text{num}})^n(u^0) \]

We want that the ratio \(\frac{\|u^n\|}{\|u^0\|}\) stay bounded for all \(u^0\) and for all \((n \Delta t) \leq T\).

This is equivalent to that the norm \(\| (L_{\text{num}})^n \| \) remain bounded for all \((n \Delta t) \leq T\), which leads to the definition of stability below.
Definition: (numerical stability of linear methods)

Consider numerical method, \( u^{n+1} = L_{\text{num}}(u^n) \).

Suppose for any \( T > 0 \), there exists a constant \( C_T \) such that

\[
\left\| \left( L_{\text{num}} \right)^n \right\| \leq C_T \quad \text{for all } n\Delta t \leq T
\]

Then we say the numerical method \( L_{\text{num}} \) is stable.

Theorem:

If \( \left\| L_{\text{num}} \right\| \leq 1 + C \Delta t \), then the method \( L_{\text{num}} \) is stable.

Proof:

\[
\left\| \left( L_{\text{num}} \right)^n \right\| \leq \left( \left\| L_{\text{num}} \right\| \right)^n \leq (1 + C \Delta t)^n \leq \exp \left( C n \Delta t \right) = \exp \left( C T \right) \quad \text{for all } n\Delta t \leq T
\]

Remark: \( \left\| L_{\text{num}} \right\| \leq 1 + C \Delta t \) is easier to check than \( \left\| (L_{\text{num}})^n \right\| \leq C_T \).

Stability of the FTCS method (IVP)

\[
\begin{align*}
  u_i^{n+1} &= u_i^n + r \left( u_{i+1}^n - 2u_i^n + u_{i-1}^n \right), \quad -\infty < i < +\infty, \quad r = \frac{\Delta t}{(\Delta x)^2}
\end{align*}
\]

We write the method as

\[
  u_i^{n+1} = ru_{i+1}^n + (1 - 2r)u_i^n + ru_{i-1}^n
\]

We consider the infinity norm:

\[
\| \bar{u} \| = \sup_{-\infty < i < +\infty} | u_i |
\]

We fix \( r = \frac{\Delta t}{(\Delta x)^2} \) while \( \Delta x \to 0 \) and \( \Delta t \to 0 \).

Theorem:

\[
\begin{align*}
  \text{FTCS is} & \quad \begin{cases} 
    \text{stable} & \text{if } r \leq \frac{1}{2} \\
    \text{unstable} & \text{if } r > \frac{1}{2}
  \end{cases}
\end{align*}
\]

Proof:
For \( r \leq \frac{1}{2} \), all three coefficients, \( r, (1-2r) \) and \( r \), are non-negative.

\[
|u_{i+1}^n| \leq r |u_{i+1}^n| + (1 - 2r) |u_i^n| + r |u_{i-1}^n|
\leq r \|u^n\| + (1 - 2r) \|u^n\| + r \|u^n\| = \|u^n\|
\]

Since this is true at all values of \( i \), we conclude

\[
\|u^{n+1}\| \leq \|u^n\| \quad \text{for all } u^n
\]

\Rightarrow \quad \text{It is stable.}

For \( r > \frac{1}{2} \), we notice that coefficients \( r, (1-2r) \), and \( r \) alternate in sign.

\[ r > 0 \quad , \quad (1-2r) < 0 \quad , \quad r > 0 \]

Consider a solution of the special form:

\[ u_i^n = \rho^n (-1)^i \]

where \( \rho \) is called the magnification factor.

Substituting into the FTCS method, we obtain

\[
\rho^{n+1} (-1)^i = r \rho^n (-1)^{i+1} + (1-2r) \rho^n (-1)^i + r \rho^n (-1)^i - 1
\]

\Rightarrow \quad \rho = r + (1 - 2r) - r = 1 - 4r

For \( r > 1/2 \), the magnification factor \( \rho = (1-4r) \) satisfies

\[
|\rho| = 4r - 1 > 1
\]

\Rightarrow \quad u_i^n = \rho^n (-1)^i \quad \text{grows unbounded as } n \to \infty.

\Rightarrow \quad \text{It is unstable.}

End of proof

Stability of the BTCS method (IVP)

\[
 u_{i+1}^{n+1} = u_i^n + r \left(u_{i+1}^{n+1} - 2u_i^{n+1} + u_{i-1}^{n+1}\right), \quad -\infty < i < +\infty, \quad r = \frac{\Delta t}{(\Delta x)^2} \quad \text{(E01)}
\]

Due to the infinite domain of IVP, the BTCS method (E01) needs an additional constraint to ensure that \( u^{n+1} \) is uniquely defined given \( u^n \).

At each fixed \( n \), \( |u_i^{n+1}| \) is bounded as \( |i| \to \infty \) \quad \text{(C01)}

Claim: without constraint (C01), \( u^{n+1} \) is not unique!
Proof:

Consider the special case of \( \{ u_i^n \equiv 0 , \ -\infty < i < \infty \} \).

It is straightforward to verify that \( \{ u_i^{n+1} = 0 , \ -\infty < i < \infty \} \) is a solution of (E01).

Besides this trivial solution, we look for a solution of the form \( u_i^{n+1} = \alpha^i \).

Substituting \( u_i^n \equiv 0 \) and \( u_i^{n+1} = \alpha^i \) into (E01), we get

\[
\alpha^i = 0 + r \alpha^i \left( \alpha - 2 + \frac{1}{\alpha} \right)
\]

\[<===>\]
\[
\alpha - 2 + \frac{1}{\alpha} = \frac{1}{r}
\]

\[<===>\]
\[
\alpha^2 - \left( 2 + \frac{1}{r} \right) \alpha + 1 = 0
\]

The quadratic equation has two roots: \(|\alpha_1| > 1\) and \(|\alpha_2| < 1\).

\( u_i^{n+1} = (\alpha_1)^i \) is a solution of (E01) satisfying \( \lim_{i \to \infty} |u_i^{n+1}| = \infty \).

\( u_i^{n+1} = (\alpha_2)^i \) is a solution of (E01) satisfying \( \lim_{i \to -\infty} |u_i^{n+1}| = \infty \).

Therefore, to uniquely define \( u^{n+1} \) in (E01), we need to impose (C01).

Theorem:

The BTCS method is unconditionally stable. That is, it is stable for any \( r > 0 \).

Proof:

We rewrite the BTCS as

\[
u_i^n = (1 + 2r)u_i^{n+1} - r(u_{i+1}^{n+1} + u_{i-1}^{n+1})\]

For any \( \varepsilon > 0 \), there exists \( i_0 \) such that

\[
\sup_i |u_i^{n+1}| \geq |u_{i_0}^{n+1}| \geq \left( \sup_i |u_i^{n+1}| - \varepsilon \right)
\]

At index \( i_0 \), the BTCS gives us

\[
|u_{i_0}^n| \geq (1 + 2r) |u_{i_0}^{n+1}| - r |u_{i_0+1}^{n+1} + u_{i_0-1}^{n+1}|
\]

\[\geq (1 + 2r) \left( \sup_i |u_i^{n+1}| - \varepsilon \right) - 2r \cdot \sup_i |u_i^{n+1}|\]
\[ \|u^n\| \geq \sup_i \|u^{n+1}_i\| - (1+2r)\varepsilon \]
\[ \sup_i \|u^n_i\| \geq \sup_i \|u^{n+1}_i\| - (1+2r)\varepsilon \]

Since this is true for any \( \varepsilon > 0 \), we conclude
\[ \sup_i \|u^n_i\| \geq \sup_i \|u^{n+1}_i\| \]

That is,
\[ \|u^{n+1}\|_\infty \leq \|u^n\|_\infty \quad \text{for all } u^n \]

\[ \Rightarrow \text{It is stable for any } r > 0. \]

**Function norm, vector norm, norm of numerical solution**

We clarify the definition of norm for numerical solution.

We first look at the new situation we are facing in numerical solutions of PDEs, in comparison with numerical solutions of ODEs.

**Numerical solution at time level \( n \):**

- **ODE:** \( u_n \approx u(t_n), \quad \text{a vector of fixed size;} \)
  
  \[ \text{size does not increase with numerical resolution.} \]

- **PDE:** \( u^n = \left\{ u^n_i, \ 1 \leq i \leq N \right\} = \left\{ u(x_i, t_n), \ 1 \leq i \leq N \right\}, \quad \text{a vector of size } N; \)

  \[ \text{size increases with numerical resolution.} \]

For a PDE, the numerical solution at \( t_n \) is both

- a discrete vector and
- an approximation to a continuous function.

This new situation demands that the norm of numerical solution \( u^n \) should have both features of the vector norm and those of the function norm.

Consider a continuous function \( u(x) \) over \([a, b]\) and a discrete version of \( u(x)\):

\[ u(x), \quad a \leq x \leq b \]

\[ \tilde{u} = \left\{ u_i = u(x_i), \ 1 \leq i \leq N \right\} \]

**Function norm:**

\[ \|u\|_p = \left( \int_a^b |u(x)|^p \, dx \right)^{\frac{1}{p}} = \text{finite} \]
Vector norm:

\[ \| \tilde{u} \|_p = \left( \sum_{i=1}^{N} |u_i|^p \right)^{\frac{1}{p}} \]

\[ = \left( \frac{1}{\Delta x} \sum_{i=1}^{N} |u(x_i)|^p \Delta x \right)^{\frac{1}{p}} = \left( \frac{1}{\Delta x} \right)^{\frac{1}{p}} \left( \sum_{i=1}^{N} |u(x_i)|^p \Delta x \right)^{\frac{1}{p}} \]

\[ \approx \left( \frac{1}{\Delta x} \right)^{\frac{1}{p}} \left( \int_{a}^{b} |u(x)|^p \, dx \right)^{\frac{1}{p}} \to \infty \quad \text{as } \Delta x \to 0 \]

To ensure that the norm of numerical solution converge as \( \Delta x \to 0 \), we adopt the norm below for numerical solutions of PDEs.

**Norm of numerical solution:**

\[ \| \tilde{u} \|_p = \left( \sum_{i=1}^{N} |u_i|^p \Delta x \right)^{\frac{1}{p}} \]