

Financial Mathematics 32000

Lecture 3

Roger Lee

2023 April 5

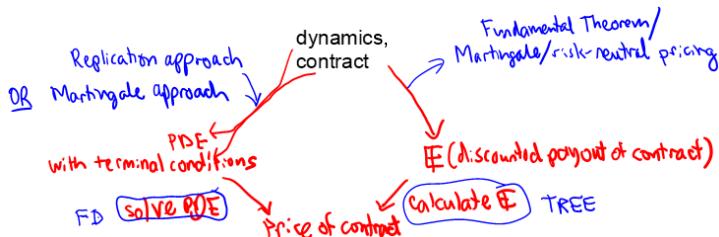
UNIT 2: Finite Difference Methods

Let $C(S, t, \dots)$ be the price of the derivative contract as a function of the underlying S , time t , etc. Find PDE satisfied by C .

Finite difference methods solve a discretized version of that PDE:

They compute an approximation of C at each node (S, t) in a grid.

The computation uses a discretization of the PDE, relating the value at each node to the values at nearby nodes.



Explicit FD scheme

Error analysis: first approach

Error analysis: second approach

Second approach, applied to explicit FD scheme

Transformed Black-Scholes PDE

Find the price C of an option that pays $f(S)$ at time T , assuming

$$dS_t = R_{grow} S_t dt + \sigma S_t dW_t$$

$$d \log S_t = \nu dt + \sigma dW_t \quad \nu := R_{grow} - \sigma^2/2$$

In original variables: $C(S, t)$ has drift rC , so

$$\frac{\partial C}{\partial t} + r \frac{\partial C}{\partial S} + \frac{1}{2} \sigma^2 \frac{\partial^2 C}{\partial S^2} = rC$$

$$C(S, T) = f(S)$$

In changed variables $x = \log S$, likewise $C(x, t)$ has drift rC , so

$$\frac{\partial C}{\partial t} + \nu \frac{\partial C}{\partial x} + \frac{1}{2} \sigma^2 \frac{\partial^2 C}{\partial x^2} = rC$$

$$C(x, T) = f(e^x)$$

Finite difference grid

- ▶ Now introduce a grid with N time intervals of length Δt and $2J$ space intervals of length Δx .

Let $\Delta t = T/N$. More to say about Δx later.

Let $t_n = n\Delta t$ for $n = 0, \dots, N$.

Let $x_j = x_0 + j\Delta x$ for $j = -J, \dots, J$.

- ▶ We want to approximate the true PDE solution at (x_j, t_n) :

$$\text{Exact:} \quad C(x_j, t_n)$$

using a finite difference solution

$$\text{Approximate:} \quad C_n^j$$

which will be determined according to a finite difference scheme.

Finite difference grid

Explicit FD scheme

○○●○○○

Error analysis: first approach

○○○○○○○○○○

Error analysis: second approach

○○○○○○○

Second approach, applied to expl

○○○○○

Explicit finite difference scheme

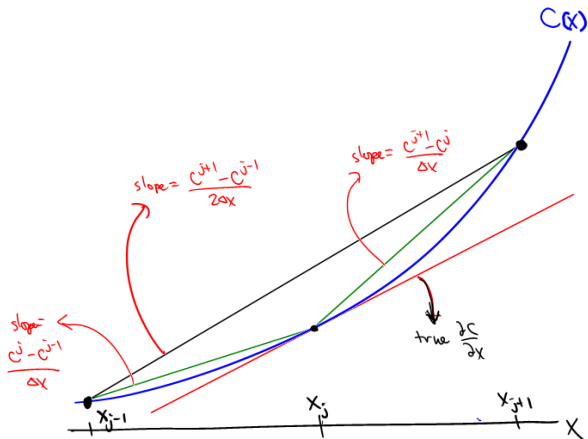
Terminal condition: $C_N^j := f(e^{x_j})$ for all j .

Boundary conditions for a *call*: $C_n^{-J} = 0$ and $C_n^J = e^{x_J} - Ke^{-r(T-t_n)}$ for all n . (Other contracts will have different boundary conditions.)

The explicit [and central-space] FD scheme approximates the PDE at (x_j, t_{n+1}) using these finite differences in place of partial derivatives:

$$\begin{aligned}\frac{\partial C}{\partial t} &\approx \frac{C_{n+1}^j - C_n^j}{\Delta t} \\ \frac{\partial C}{\partial x} &\approx \frac{C_{n+1}^{j+1} - C_{n+1}^{j-1}}{2\Delta x} \\ \frac{\partial^2 C}{\partial x^2} &\approx \frac{1}{\Delta x} \left(\frac{C_{n+1}^{j+1} - C_{n+1}^j}{\Delta x} - \frac{C_{n+1}^j - C_{n+1}^{j-1}}{\Delta x} \right) = \frac{C_{n+1}^{j+1} - 2C_{n+1}^j + C_{n+1}^{j-1}}{(\Delta x)^2}\end{aligned}$$

One-sided and central finite differences



Finite difference equation

- Then the finite difference equation is

$$\frac{C_{n+1}^j - C_n^j}{\Delta t} + \nu \frac{C_{n+1}^{j+1} - C_{n+1}^{j-1}}{2\Delta x} + \frac{1}{2}\sigma^2 \frac{C_{n+1}^{j+1} - 2C_{n+1}^j + C_{n+1}^{j-1}}{(\Delta x)^2} = rC_{n+1}^j$$

Let's replace the rC_{n+1}^j by rC_n^j ; this will have no effect on rate of convergence.

- Now solve for C_n^j :

$$C_n^j = \frac{1}{1 + r\Delta t} (q_u C_{n+1}^{j+1} + q_m C_{n+1}^j + q_d C_{n+1}^{j-1})$$

where ...

Finite difference solution

... where

$$\begin{aligned}
 q_u &= \frac{1}{2} \left[\frac{\sigma^2 \Delta t}{(\Delta x)^2} + \frac{\nu \Delta t}{\Delta x} \right] &= p_u - \frac{\nu^2 (\Delta t)^2}{2(\Delta x)^2} \\
 q_m &= 1 - \frac{\sigma^2 \Delta t}{(\Delta x)^2} &= p_m + \frac{\nu^2 (\Delta t)^2}{(\Delta x)^2} \\
 q_d &= \frac{1}{2} \left[\frac{\sigma^2 \Delta t}{(\Delta x)^2} - \frac{\nu \Delta t}{\Delta x} \right] &= p_d - \frac{\nu^2 (\Delta t)^2}{2(\Delta x)^2}
 \end{aligned}$$

- ▶ Then induct backwards from $n = N - 1$ to $n = 0$, obtaining C at all nodes in the grid.
- ▶ Hence the explicit FD scheme is “equivalent” to a trinomial tree. It has $1/(1 + r\Delta t)$ instead of $e^{-r\Delta t}$, and $q_{u,m,d}$ instead of $p_{u,m,d}$, but these changes have no effect on convergence rate.

Explicit FD scheme

Error analysis: first approach

Error analysis: second approach

Second approach, applied to explicit FD scheme

Definitions of error

Approximation error is defined as (FD solution) - (true PDE solution):

$$\mathbf{e}_n^j := \mathbf{C}_n^j - C(x_j, t_n)$$

Local discretization error or *truncation error* measures how well the PDE solution satisfies the L3.9 FD equation: plug the PDE solution into each side of the FD equation, and take the difference LHS – RHS:

$$\begin{aligned} L(x, t) := & \frac{C(x, t+\Delta t) - C(x, t)}{\Delta t} + \nu \frac{C(x+\Delta x, t+\Delta t) - C(x-\Delta x, t+\Delta t)}{2\Delta x} \\ & + \frac{\sigma^2}{2} \frac{C(x+\Delta x, t+\Delta t) - 2C(x, t+\Delta t) + C(x-\Delta x, t+\Delta t)}{(\Delta x)^2} - rC(x, t) \end{aligned}$$

$$\text{Note } \frac{(\Delta t)L(x, t)}{1+r\Delta t} = \frac{q_u C(x+\Delta x, t+\Delta t) + q_m C(x, t+\Delta t) + q_d C(x-\Delta x, t+\Delta t)}{1+r\Delta t} - C(x, t)$$

= error in approximating $C(x, t)$, if $C(\cdot, t + \Delta t)$ are known exactly.

So $(\Delta t)L$ = error *newly* introduced in column n , if column $n + 1$ exact.

Size of local discretization (truncation) error

How big is L ? Taylor expansion of PDE solution C about $(x, t + \Delta t)$:

$$C(x, t) = C - (\Delta t)C_t + O(\Delta t)^2$$

$$C(x \pm \Delta x, t + \Delta t) = C \pm (\Delta x)C_x + \frac{(\Delta x)^2}{2}C_{xx} \pm \frac{(\Delta x)^3}{6}C_{xxx} + O(\Delta x)^4$$

where on the RHS, C and its partials are evaluated at $(x, t + \Delta t)$. So

$$\begin{aligned} L(x, t) &= \frac{(\Delta t)C_t + O(\Delta t)^2}{\Delta t} + \nu \frac{2(\Delta x)C_x + O(\Delta x)^3}{2\Delta x} + \frac{\sigma^2}{2} \frac{(\Delta x)^2 C_{xx} + O(\Delta x)^4}{(\Delta x)^2} - rC \\ &= C_t + O(\Delta t) + \nu(C_x + O(\Delta x)^2) + \frac{\sigma^2}{2}(C_{xx} + O(\Delta x)^2) - rC \\ &= O(\Delta t) + O(\Delta x)^2 \end{aligned}$$

For all (x, t) we have $L(x, t) \rightarrow 0$ as $\Delta t, \Delta x \rightarrow 0$, so we say the explicit FD scheme is *consistent* (to first order in Δt and second order in Δx)

Size of approximation error

Approximation error \mathbf{e}_n^j combines the error *newly* introduced in n ,
and propagation of *existing* error from $n + 1$. Write $\mathbf{L}_n^j := L(x_j, t_n)$.

$$\begin{aligned}\mathbf{e}_n^j &:= \mathbf{C}_n^j - C(x_j, t_n) \\ &= \frac{1}{1 + r\Delta t} (q_u \mathbf{C}_{n+1}^{j+1} + q_m \mathbf{C}_{n+1}^j + q_d \mathbf{C}_{n+1}^{j-1}) \\ &\quad - \frac{1}{1 + r\Delta t} (q_u C(x_{j+1}, t_{n+1}) + q_m C(x_j, t_{n+1}) + q_d C(x_{j-1}, t_{n+1}) - (\Delta t) \mathbf{L}_n^j) \\ &= \frac{1}{1 + r\Delta t} \left(\underbrace{q_u \mathbf{e}_{n+1}^{j+1} + q_m \mathbf{e}_{n+1}^j + q_d \mathbf{e}_{n+1}^{j-1}}_{\text{error propagated from column } n+1 \text{ to } n} + \underbrace{(\Delta t) \mathbf{L}_n^j}_{\text{new error in column } n} \right)\end{aligned}$$

Let $\mathbf{E}_n := \max_j |\mathbf{e}_n^j|$. If $r \geq 0$ then

$$\mathbf{E}_n \leq (|q_u| + |q_m| + |q_d|) \mathbf{E}_{n+1} + (\Delta t) \bar{\mathbf{L}}$$

where $\bar{\mathbf{L}} := \max_{j,n} |\mathbf{L}_n^j|$.

Approximation error goes to zero, assuming stability

If we have the *stability* condition $q_u, q_m, q_d \geq 0$, then we can drop the absolute values and get

$$E_n \leq E_{n+1} + (\Delta t)\bar{L}.$$

Terminal approximation error is $E_N = 0$, so max approximation error

$$\max_{0 \leq n \leq N} E_n \leq N(\Delta t)\bar{L} = T\bar{L} \longrightarrow 0 \quad \text{as } \Delta x, \Delta t \rightarrow 0$$

(because $L \rightarrow 0$ by the consistency analysis).

Hence we say the explicit FD scheme is *convergent*.

The rate of this convergence is $O(\Delta t) + O(\Delta x)^2$.

General principle: consistency plus stability implies convergence.

The stability condition: comments

The stability condition is that the weights be nonnegative:

$$\begin{aligned}
 q_u, q_d \geq 0 & \Leftrightarrow \frac{\sigma^2 \Delta t}{(\Delta x)^2} \pm \frac{\nu \Delta t}{\Delta x} \geq 0 & \Leftrightarrow & \Delta x \leq \frac{\sigma^2}{|\nu|} \\
 q_m \geq 0 & \Leftrightarrow 1 - \frac{\sigma^2 \Delta t}{(\Delta x)^2} \geq 0 & \Leftrightarrow & \Delta t \leq \frac{1}{\sigma^2} (\Delta x)^2
 \end{aligned}$$

Intuition: Think in terms of trinomial tree.

- ▶ We want to assign weights on the branches to produce variance $\sigma^2 \Delta t$.
- ▶ But if $\sigma^2 \Delta t > (\Delta x)^2$, then we are trying to squeeze too much variance into too small a structure, resulting in wing weights that sum to more than 100%, which can cause errors to blow up.

The stability condition: comments

- ▶ The condition $\Delta x \leq \sigma^2/|\nu|$ is usually not burdensome, but can be if the drift is strong.
- ▶ The condition $\Delta t \leq (\Delta x)^2/\sigma^2$ can be burdensome. It could force you to choose smaller (hence more) time steps than accuracy alone would dictate. So we want to look for alternative FD schemes not subject to this stability constraint.
- ▶ In the constant-volatility trinomial tree, the recommendation was $\Delta x = \sigma\sqrt{3\Delta t}$, which yields $\Delta t = (\Delta x)^2/(3\sigma^2) < (\Delta x)^2/\sigma^2$, satisfying the stability condition.

The stability condition: comments

- If we had worked with a PDE in S instead of $x = \log S$,

$$\frac{\partial \tilde{C}}{\partial t} + rS \frac{\partial \tilde{C}}{\partial S} + \frac{1}{2} \sigma^2 S^2 \frac{\partial^2 \tilde{C}}{\partial S^2} = r\tilde{C}, \quad \tilde{C}(S, t) = C(\log S, t)$$

then the stability constraint becomes $\Delta t \leq (\Delta S)^2 / (\sigma^2 S^2)$, which is even more wasteful of resources, because a subset of the grid (the high- S part) forces us to choose Δt very small – even though the rest of the grid does not need such small Δt .

- But we may want to do a grid for S , not $\log S$, for reasons such as accuracy in modelling payoffs that are piecewise linear in S (not $\log S$). This makes it even more important to find FD schemes which do not have stability restrictions.

Consistency: comments

Take $\nu = r = 0$. Recall the local discretization (truncation) error

$$L(x, t) := \frac{C(x, t + \Delta t) - C(x, t)}{\Delta t} + \frac{\sigma^2}{2} \frac{C(x + \Delta x, t + \Delta t) - 2C(x, t + \Delta t) + C(x - \Delta x, t + \Delta t)}{(\Delta x)^2}$$

Taylor expansion of PDE solution C about $(x, t + \Delta t)$:

$$\begin{aligned} C(x, t) &= C - (\Delta t)C_t + \frac{(\Delta t)^2}{2}C_{tt} + O(\Delta t)^3 \\ C(x \pm \Delta x, t + \Delta t) &= C \pm (\Delta x)C_x + \frac{(\Delta x)^2}{2}C_{xx} \pm \frac{(\Delta x)^3}{6}C_{xxx} \\ &\quad + \frac{(\Delta x)^4}{24}C_{xxxx} \pm \frac{(\Delta x)^5}{120}C_{xxxxx} + O(\Delta x)^6 \end{aligned}$$

where on the RHS, C and its partials are evaluated at (x, t) . So

$$\begin{aligned} L(x, t) &= \frac{(\Delta t)C_t - (1/2)(\Delta t)^2C_{tt} + O(\Delta t)^3}{\Delta t} + \frac{\sigma^2}{2} \frac{(\Delta x)^2C_{xx} + (1/12)(\Delta x)^4C_{xxxx} + O(\Delta x)^6}{(\Delta x)^2} \\ &= C_t - \frac{\Delta t}{2}C_{tt} + O(\Delta t)^2 + \frac{\sigma^2}{2} \left(C_{xx} + \frac{(\Delta x)^2}{12}C_{xxxx} + O(\Delta x)^4 \right) \end{aligned}$$

Consistency: comments

Then

$$\begin{aligned} L(x, t) &= \frac{\Delta t}{2} C_{tt} + \frac{\sigma^2 (\Delta x)^2}{2 \cdot 12} C_{xxxx} + O(\Delta t)^2 + O(\Delta x)^4 \\ &= O(\Delta t) + O(\Delta x)^2 \longrightarrow 0 \quad \text{as } \Delta t, \Delta x \rightarrow 0 \end{aligned}$$

so the explicit FD scheme is accurate to first order in Δt and second order in Δx . But maybe we can do better:

$$C_{tt} = \left(-\frac{1}{2} \sigma^2 C_{xx} \right)_t = -\frac{1}{2} \sigma^2 C_{txx} = \left(\frac{1}{2} \sigma^2 \right)^2 C_{xxxx}$$

So

$$-\frac{\Delta t}{2} C_{tt} + \frac{\sigma^2 (\Delta x)^2}{2 \cdot 12} C_{xxxx} = \frac{\sigma^2}{2} \left(\frac{(\Delta x)^2}{12} - \frac{\Delta t}{2} \frac{\sigma^2}{2} \right) C_{xxxx}$$

To make **leading order terms** vanish:

Caveats

Some caveats:

- ▶ We've assumed that boundary conditions have no error.
- ▶ We've assumed that L is *uniformly* $O(\Delta t) + O(\Delta x)^2$ in the entire domain. This is true if C_{tt} and C_{xxxx} are bounded on the domain – which does not hold for discontinuous or kinked payoffs.
- ▶ We've assumed constant σ

Explicit FD scheme

Error analysis: first approach

Error analysis: second approach

Second approach, applied to explicit FD scheme

Error analysis of general FD schemes

We want to analyze error of general FD schemes, not just the explicit scheme. General FD schemes are more complicated, so we will make some changes to facilitate the analysis.

- ▶ Analyze the “model” problem, which is the heat equation (to which the constant-coefficients B-S PDE can be transformed).
- ▶ Introduce general notions of consistency, stability, convergence.
- ▶ Use “Fourier” method to do stability analysis (as opposed to the “maximum principle” analysis of the previous section).

The model problem

Recall the B-S PDE can be transformed into the “model” problem

$$U_\tau = \frac{1}{2}\sigma^2 U_{xx}$$

with initial conditions given at $\tau = 0$.

This problem is easier to study.

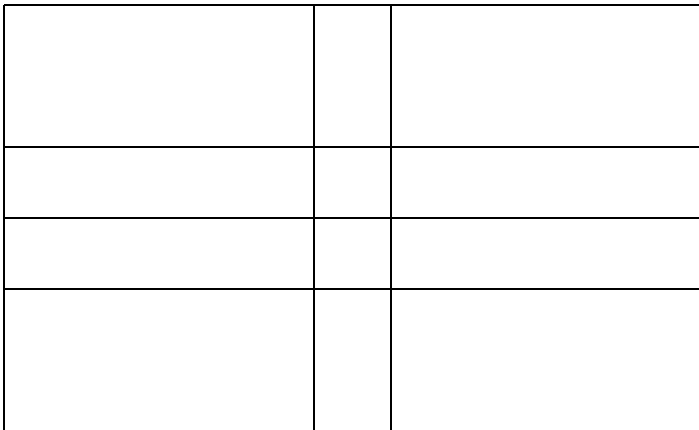
Let $\Delta t = T/N$ and $\tau_n = n\Delta t$ and $x_j = x_0 + j\Delta x$ for $j = -J, \dots, J$.

For example, the explicit FD scheme becomes

$$\frac{U_{n+1}^j - U_n^j}{\Delta t} - \frac{1}{2}\sigma^2 \frac{U_n^{j+1} - 2U_n^j + U_n^{j-1}}{(\Delta x)^2} = 0$$

where U_n^j is intended to approximate $U(x_j, \tau_n)$.

Finite difference grid



Explicit FD scheme

○○○○○○○○

Error analysis: first approach

○○○○○○○○○○

Error analysis: second approach

○○○●○○○

Second approach, applied to expl

○○○○○

FD operator

Consider a finite difference scheme

$$\mathbf{U}_{n+1} = Q(\mathbf{U}_n)$$

where the vector $\mathbf{U}_n \in \mathbb{R}^{2J+1}$ is column n of the grid, and Q is the finite difference operator that produces column $n + 1$ given column n . For some types of boundary conditions, such as zero conditions, Q will be linear operator.

Consistency

- ▶ In column n of the grid, define the local discretization (truncation) error $\mathbf{L}_n \in \mathbb{R}^{2J+1}$ by

$$\mathbf{u}_{n+1} = Q\mathbf{u}_n + (\Delta t)\mathbf{L}_n$$

where $\mathbf{u}_n \in \mathbb{R}^{2J+1}$ is the actual PDE solution sampled at the points (x_n, τ_j) for $j = -J, \dots, J$.

- ▶ Let $\bar{\mathbf{L}} := \max_n \|\mathbf{L}_n\|$. The max is over all n such that $n\Delta t \leq T$, and $\|\cdot\|$ is some vector norm.
- ▶ If $\bar{\mathbf{L}} \rightarrow 0$ as $\Delta x, \Delta t \rightarrow 0$, then we say the FD scheme is *consistent*.

Stability

- ▶ A FD scheme is *unconditionally stable* if there exists a constant A such that for all “initial data” $\mathbf{V}_0 \in \mathbb{R}^{2J+1}$, and all Δx and Δt , we have

$$\max_{n: n\Delta t \leq T} \|\mathbf{V}_n\| \leq A\|\mathbf{V}_0\|$$

where $\mathbf{V}_n := Q^n \mathbf{V}_0$ is the result of n applications of Q to \mathbf{V}_0 .

- ▶ If we need restrictions on Δx and Δt for this to hold, then the scheme is *conditionally stable*.

Convergence

- *Approximation error* is $\mathbf{e}_n := \mathbf{U}_n - \mathbf{u}_n \in \mathbb{R}^{2J+1}$
- If

$$\max_{n: n\Delta t \leq T} \|\mathbf{e}_n\| \longrightarrow 0 \quad \text{as } \Delta x, \Delta t \rightarrow 0$$

then we say the FD scheme is *convergent*.

- Consistency plus stability implies convergence.

This is part of the Lax Equivalence Theorem.

Idea of proof: At each time slice, Q introduces error $\leq (\Delta t)\bar{L}$.

Error propagates to the end, becoming $\leq A(\Delta t)\bar{L}$ by stability.

But error was introduced at each of N time slices, so

$$\text{Total approximation error} \leq AN(\Delta t)\bar{L} = AT\bar{L} \longrightarrow 0$$

by consistency.

Explicit FD scheme

Error analysis: first approach

Error analysis: second approach

Second approach, applied to explicit FD scheme

Stability, via “Fourier” analysis

Look for solutions to the FD scheme of the form

$$\mathbf{U}_n^j = \lambda_k^n e^{i\pi k j/J} \quad \text{for } k = -J+1, \dots, J.$$

where the superscripts on the RHS are powers, and $i = \sqrt{-1}$.

We call k the wave number and λ_k the amplification factor.

Think of the “Fourier modes” $\mathbf{m}_k := (e^{i\pi k j/J})_{j=-J, \dots, J} \in \mathbb{C}^{2J+1}$ as “basis vectors” for a subspace $\mathcal{P} \subset \mathbb{C}^{2J+1}$. Picture, for example, the real part of $e^{i\pi k j/J}$ for $k = 1$ and $k = J$:

Action of FD operator

What happens to these Fourier modes as we apply the FD operator?

If $|\lambda_k| \leq 1$ for all k , then these basis functions don't blow up.

Therefore, general initial data \mathbf{V}_0 won't blow up. So we have stability.

Specifically, for *any* $\mathbf{V}_0 \in \mathcal{P}$, there exist coefficients c_k with

$$\mathbf{V}_0 = \sum_{k=-J+1}^J c_k \mathbf{m}_k$$

After n applications of Q to each side,

$$\mathbf{V}_n = \sum_{k=-J+1}^J c_k \lambda_k^n \mathbf{m}_k$$

So if $|\lambda_k| \leq 1$ for all k then $\|\mathbf{V}_n\| \leq \|\mathbf{V}_0\|$, hence stability (in Euclidean norm).

Application to explicit FD scheme

Substitute $U_n^j = \lambda_k^n e^{i\pi k j/J}$ into the explicit FD scheme

$$\frac{U_{n+1}^j - U_n^j}{\Delta t} - \frac{1}{2}\sigma^2 \frac{U_{n+1}^{j+1} - 2U_n^j + U_{n+1}^{j-1}}{(\Delta x)^2} = 0$$

to obtain

$$\frac{\lambda_k^{(n+1)} e^{i\pi k j/J} - \lambda_k^n e^{i\pi k j/J}}{\Delta t} = \frac{1}{2}\sigma^2 \frac{\lambda_k^n e^{i\pi k (j+1)/J} - 2\lambda_k^n e^{i\pi k j/J} + \lambda_k^n e^{i\pi k (j-1)/J}}{(\Delta x)^2}$$

Cancel U_n^j from both sides and let $\alpha := \frac{1}{2}\sigma^2 \frac{\Delta t}{(\Delta x)^2}$. Then

$$\lambda_k - 1 = \alpha(e^{i\pi k/J} - 2 + e^{-i\pi k/J}) = \alpha\left(2\cos\frac{\pi k}{J} - 2\right) = -4\alpha\sin^2\frac{\pi k}{2J}$$

because $\cos 2\theta = 1 - 2\sin^2\theta$. So

$$\lambda_k = 1 - 4\alpha\sin^2\frac{\pi k}{2J}$$

Stability condition

Note that $\sin^2 \frac{\pi k}{2J} \leq 1$, with equality iff $k = J$.

- So if $\alpha \leq 1/2$, then

$$|\lambda_k| = \left| 1 - 4\alpha \sin^2 \frac{\pi k}{2J} \right| \leq 1 \quad \text{for all } k.$$

- If $\alpha > 1/2$, then the high-frequency modes ($k \approx J$) will have $|\lambda_k| > 1$ and blow up.

The stability condition $\alpha \leq 1/2$ is equivalent to $\Delta t \leq (\Delta x)^2/\sigma^2$, the same condition as we found in the L3.16 stability analysis.