Numerical ~~Ex~~ Exercises 1

1. A ~~[scribble]~~ partial differential equation like the given equation

$$u_t - a(t)u_x - \varepsilon u_{xx} = 0 \,,$$ is linear if the

dependent variable ~~in this case x and t~~ and it's partial derivatives are present linearly in the equation. For example, for a second order linear PDE like we have here, it must take the form

$$S_1(x,t)u_{xx} + S_2(x,t)u_{xt} + S_3(x,t)u_{tx} + S_4(x,t)u_{tt} \cdots$$

$$+ S_5(x,t)u_x + S_6(x,t)u_t + S_7(x,t)u = f(x,t) \,.$$

Clearly this is satisfied with $S_6 = 1$, $S_5 = -a(t)$, $S_1 = -\varepsilon$ and $S_2 = S_3 = S_4 = S_7 = f = 0$.

The equation includes the diffusive term '$-\varepsilon u_{xx}$' and describes how the dependent variable is distributed in a given region over time. There is an additional advective ~~time~~ term '$-a(t)u_x$' which describes the motion of u as it is advected by the velocity field.

2. The θ-scheme is given by

$$U_j^{n+1} - U_j^n = \mu\left[\theta \delta_x^2 U_j^{n+1} + (1-\theta)\delta_x^2 U_j^n\right]$$

we ~~a~~ taylor expand each term around the point $(x_j, t_{n+1/2})$ for this scheme. This gives:

$$u_j^{n+1} \approx u(x, t+\tfrac{1}{2}\Delta t) + u_t\left(\tfrac{1}{2}\Delta t\right) + \frac{1}{2!}u_{tt}\left(\tfrac{1}{2}\Delta t\right)^2 + \frac{1}{3!}u_{ttt}\left(\tfrac{1}{2}\Delta t\right)^3 + \dots$$

This follows from the form of the Taylor series

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \dots$$

where (in 2 independent variables) we take $a = (x, t+\frac{1}{2}\Delta t)$ and $x = (x, t+\Delta t)$. Similarly,

$$u_j^{n} \approx u(x, t+\tfrac{1}{2}\Delta t) + u_t\left(-\tfrac{1}{2}\Delta t\right) + \frac{1}{2!}u_{tt}\left(-\tfrac{1}{2}\Delta t\right)^2 + \frac{1}{3!}u_{ttt}\left(-\tfrac{1}{2}\Delta t\right)^3 + \dots$$

where $a = (x, t+\frac{1}{2}\Delta t)$ and $x = (x, t)$. Hence

$$u_j^{n+1} - u_j^{n} = u(x, t+\tfrac{1}{2}\Delta t) - u(x, t-\tfrac{1}{2}\Delta t) + u_t\left(\tfrac{1}{2}\Delta t\right) - -u_t\left(\tfrac{1}{2}\Delta t\right) \dots$$

$$+ \frac{1}{2}u_{tt}\left(\tfrac{1}{4}\Delta t^2\right) \cancel{-} \frac{1}{2}u_{tt}\left(\tfrac{1}{4}\Delta t^2\right) + \dots$$

$$= u_t \Delta t + \frac{1}{24}(\Delta t)^3 u_{ttt} + \dots O\left((\Delta t)^5\right).$$

We can also Taylor expand $\delta_x^2 u_j^{n+1}$. Since

$$\delta_x^2 u(x,t) = u(x+\Delta x, t) - 2u(x,t) + u(x-\Delta x, t)$$

we again expand around point $(x_j, t_{n+\frac{1}{2}})$ and therefore $a = (x, t+\frac{1}{2}\Delta t)$ in the expansion, giving:

$$u(x+\Delta x, t+\Delta t) \approx u(x, t+\tfrac{1}{2}\Delta t) + u_x \Delta x + \frac{1}{2!}u_{xx}(\Delta x)^2 + \dots$$

~~$u(x,t) \approx u(x, t+\frac{1}{2}\Delta t) + u_x$~~

$$u(x-\Delta x, t+\Delta t) \approx u(x, t+\tfrac{1}{2}\Delta t) + u_x(-\Delta x) + \frac{1}{2!}u_{xx}(-\Delta x)^2 + \dots$$

Therefore we have

$$\delta_x^2 u(x, t+\Delta t) \approx \Big[u_{xx}(\Delta x)^2 + \frac{1}{12}(\Delta x)^4 u_{xxxx} + \dots \Big]_j^{n+1}.$$

We would get the same expansion for $\delta_x^2 u_j^n$ except the time component would be n and not $n+1$.

We now expand each term in the two expressions for $\delta_x^2 u_j^{n+1}$ and $\delta_x^2 u_j^n$ above around the point $(x_j, t_{n+1/2})$ but in powers of Δt. Therefore

$$\delta_x^2 u_j^{n+1} = \left[(\Delta x)^2 \frac{\partial^2 u(x, t+\frac{1}{2}\Delta t)}{\partial x^2} + \frac{1}{12}(\Delta x)^4 \frac{\partial^4 u(x, t+\frac{1}{2}\Delta t)}{\partial x^4} + \dots \right]$$

$$+ \frac{1}{2}\Delta t \left[(\Delta x)^2 u_{xxt} + \frac{1}{12}(\Delta x)^4 u_{xxxxt} + \dots \right]$$

$$+ \frac{1}{2!}\left(\frac{1}{2}\Delta t\right)^2 \left[(\Delta x)^2 u_{xxtt} + \frac{1}{12}(\Delta x)^4 u_{xxxxtt} + \dots \right] + \dots$$

The expansion for $\delta_x^2 u_j^n$ would be similar except the ~~coefficient~~ powers of Δt would be $t - (t + \frac{1}{2}\Delta t) = -\frac{1}{2}\Delta t$ hence giving

$$\delta_x^2 u_j^n = \left[(\Delta x)^2 u_{xx} + \frac{1}{12}(\Delta x)^4 u_{xxxx} + \dots \right]$$

$$- \frac{1}{2}\Delta t \left[(\Delta x)^2 u_{xxt} + \frac{1}{12}(\Delta x)^4 u_{xxxxt} + \dots \right]$$

$$\frac{1}{2!}\left(-\frac{1}{2}\Delta t\right)^2 \left[(\Delta x)^2 u_{xxtt} + \dots \right] - \dots$$

Combining these ~~gives~~ in the context of the θ-scheme gives:

$$\theta \delta_x^2 u_j^{n+1} + (1-\theta)\delta_x^2 u_j^n =$$

$$\left[(\Delta x)^2 u_{xx} + \frac{1}{12}(\Delta x)^4 u_{xxxx} + \dots \right]$$

$$+ \left(\theta - \frac{1}{2}\right)\Delta t \left[(\Delta x)^2 u_{xxt} + \frac{1}{12}(\Delta x)^4 u_{xxxxt} + \dots \right]$$

$$+ \frac{1}{2!}\left(-\frac{1}{2}\Delta t\right)^2 \left[(\Delta x)^2 u_{xxtt} + \frac{1}{12}(\Delta x)^4 u_{xxxxtt} + \dots \right]$$

$$+ \dots$$

using the definition of the truncation error we therefore have:

$$T_j^{n+1/2} := \frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{\theta \delta_x^2 u_j^{n+1} + (1-\theta)\delta_x^2 u_j^n}{(\Delta x)^2}$$

$$= u_t + \frac{1}{24}(\Delta t)^2 u_{ttt} + \cdots$$

$$- \left[u_{xx} + \frac{1}{12}(\Delta x)^2 u_{xxxx} + \cdots \right]$$

$$- (\theta - \tfrac{1}{2})\Delta t \left[u_{xxt} + \frac{1}{12}(\Delta x)^2 u_{xxxxt} + \cdots \right]$$

$$- \frac{1}{2!}\left(-\tfrac{1}{2}\Delta t\right)^2 \left[u_{xxtt} + \frac{1}{12}(\Delta x)^2 u_{xxxxtt} \right] + \cdots$$

$$= (u_t - u_{xx}) + \left[(\tfrac{1}{2} - \theta)\Delta t\, u_{xxt} - \frac{1}{12}(\Delta x)^2 u_{xxxx} \right]$$

$$+ \left[\frac{1}{24}(\Delta t)^2 u_{ttt} - \frac{1}{2!}\left(-\tfrac{1}{2}\Delta t\right)^2 u_{xxtt} \right]$$

$$+ \frac{1}{12}(\tfrac{1}{2} - \theta)\Delta t (\Delta x)^2 u_{xxxxt} - \cdots .$$

3. The θ-discretisation for the equation

$$u_t - a(t)u_x - \varepsilon u_{xx} = 0$$

is done as follows:

~~Using an upw~~ we can use a forward difference scheme for the time derivative appoximation, and so

$$u_t \approx \frac{u_j^{n+1} - u_j^n}{\Delta t}.$$

For the advective term an spatial upwind scheme is used in conjunction with the weight θ, therefore

$$u_x \approx \theta \frac{u_{j+1}^{n+1} - u_j^{n+1}}{\Delta x} + (1-\theta)\frac{u_{j+1}^n - u_j^n}{\Delta x}, \text{ assuming } a(t) < 0.$$

Finally, the diffusive term is weighted using the same θ as follows

$$u_{xx} \approx \theta \frac{u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}}{(\Delta x)^2} + (1-\theta)\frac{u_{j+1}^n - 2u_j^n + u_{j-1}^n}{(\Delta x)^2}.$$

Putting this together we can discretise our equation

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} - \frac{a(t)}{\Delta x}\left(\theta(u_{j+1}^{n+1} - u_j^{n+1}) + (1-\theta)(u_{j+1}^n - u_j^n)\right) \ldots$$

$$\ldots - \frac{\varepsilon}{(\Delta x)^2}\left(\theta\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right) + (1-\theta)\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right)\right)$$

$$= 0.$$

~~Defing~~ Defining two ~~constan~~ variables $k_1 = \frac{a(t)\Delta t}{\Delta x}$ and $k_2 = \frac{\varepsilon \Delta t}{(\Delta x)^2}$ we can simplify the discretisation,

assuming $a(t)$ is constant, to

$$u_j^{n+1} - u_j^n = k_1\left(\theta\left(u_{j+1}^{n+1} - u_j^{n+1}\right) + (1-\theta)\left(u_{j+1}^n - u_j^n\right) \ldots\right.$$

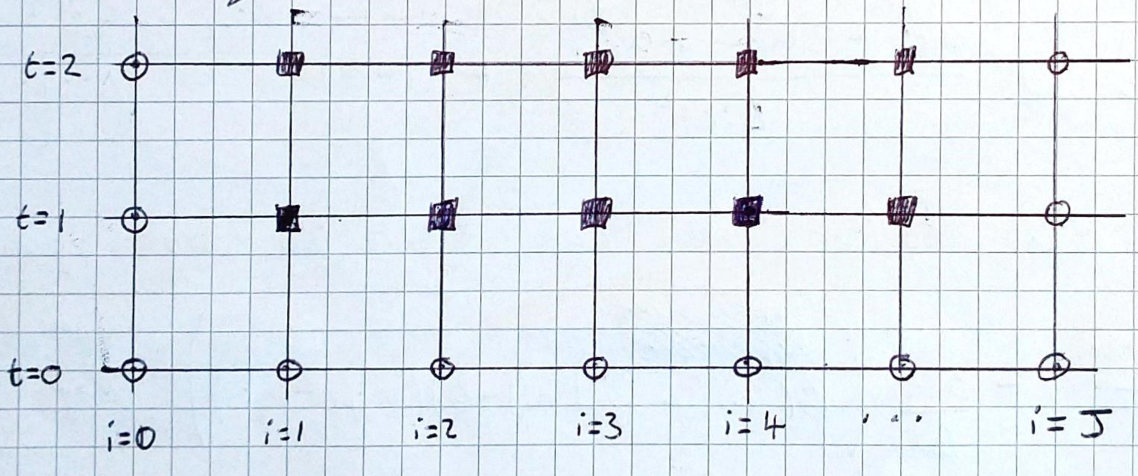
$$\left. \ldots + k_2\left(\theta\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right) + (1-\theta)\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right)\right).$$

Since all points on the ~~boundary~~ timestep $t=n$ will be known and points on the timestep $t=n+1$ will be unknown quantities to be derived we can collect known and unknown terms on either side of the discretisation, giving

$$u_{j+1}^{n+1}\left(-k_1\theta - k_2\theta\right) + u_j^{n+1}\left(1 + k_1\theta + 2k_2\theta\right) + u_{j-1}^{n+1}\left(-k_2\theta\right) = \ldots$$

$$\ldots u_{j+1}^n\left(k_1 - k_1\theta + k_2 - k_2\theta\right) + u_j^n\left(1 - k_1 + k_1\theta - 2k_2 + 2k_2\theta\right) + u_{j-1}^n\left(k_2 - k_2\theta\right).$$

If we draw a stencil for the θ-scheme at the boundary where circles denote the known boundary points and squares the interior unknowns we have



The first unknown in the scheme is indexed by $i=1$, $t=1$. The last unknown in each row (i.e. for each time step) will be indexed by $i=J-1$. Hence, the scheme can be written in the form $\underline{A}\,\underline{x} = \underline{b}$ as follows:

$$
\begin{bmatrix}
S_1 & S_2 & 0 & 0 & \cdots \\
S_3 & S_4 & S_5 & 0 & \cdots \\
0 & S_6 & S_7 & S_8 & \cdots \\
0 & 0 & S_9 & S_{10} & \cdots \\
\vdots & \vdots & \vdots & & \ddots \\
 & & & & S_n
\end{bmatrix}
\begin{bmatrix}
u_1^{n+1} \\
u_2^{n+1} \\
\vdots \\
\vdots \\
u_{J-1}^{n+1}
\end{bmatrix}
=
\begin{bmatrix}
f(u_0^n, u_1^n, u_2^n \cdots) \\
f(u_0^n, u_1^n, u_2^n \cdots) \\
\vdots \\
\vdots
\end{bmatrix}.
$$

~~At the~~ Note that the first and last elements in the $\underline{b}$ vector will have an extra boundary condition term U_0^{n+1} and U_J^{n+1} respectively.

4. For the explicit scheme we have the discretised equation $U_j^{n+1} = U_j^n + \mu\left(U_{j+1}^n - 2U_j^n + U_{j-1}^n\right)$ for the linear diffusion equation $u_t = u_{xx}$. We can now perform a fourier analysis of the scheme. Suppose we substitute into ~~oour~~ our scheme

$U_j^n = (\lambda)^n e^{ik(j\Delta x)}$, then we also know that

$U_j^{n+1} = \lambda U_j^n$ and similarly

$U_{j+1/j-1}^n = e^{\pm ik\Delta x} U_j^n$. Substituting these ~~in~~ terms into the explicit scheme equation we have

$\lambda U_j^n = U_j^n + \mu\left(e^{ik\Delta x} U_j^n - 2U_j^n + e^{-ik\Delta x} U_j^n\right)$. Dividing by U_j^n we have

$$\lambda = 1 + \mu\left(e^{ik\Delta x} - 2 + e^{-ik\Delta x}\right),$$

$$\lambda = 1 - 2\mu\left(1 - \cos(k\Delta x)\right),$$

$\lambda = 1 - 4\mu \sin^2\left(\frac{k\Delta x}{2}\right)$. As time n increases $\lambda(k)$ will diverge so long as $|\lambda(k)| > 1$ and thus the stability condition becomes

$\left|1 - 4\mu \sin^2\left(\frac{k\Delta x}{2}\right)\right| \leq 1$. Taking the mode most likely to be unstable i.e. $k\Delta x = \pi$ we have

$|1 - 4\mu| \leq 1$, Since $\mu > 0$ we look at the negative condition $1 - 4\mu \geq -1$ and so the stability condition becomes $\mu \leq \frac{1}{2}$.

This explains the question 4 python plot since for $\Delta x = 0.05$ we have

$$\mu_{0.0013} = \frac{\Delta t}{(\Delta x)^2} = \frac{0.0013}{(0.05)^2} = 0.52 \quad \text{(i.e. unstable) and}$$

$$\mu_{0.0012} = \frac{\Delta t}{(\Delta x)^2} = \frac{0.0012}{(0.05)^2} = 0.48 \quad \text{(i.e. stable).}$$

In addition, in question 4 we can observe that the initial condition gives $U_{min} = 0$ and $U_{max} = 1$ and the graphs show that all U_j^n satisfy the maximum principle

$$0 \leq U_j^n \leq 1.$$

5. Turning our attention to the θ-scheme we can similarly perform a Fourier analysis of the advective and diffusive parts of the equation in separation. Let us first rewrite the θ-scheme discretisation without the advective term i.e. setting $a=0$:

$$\frac{u_j^{n+1} - u_j^n}{\Delta t} = \frac{\varepsilon}{(\Delta x)^2}\left(\theta\left(u_{j+1}^{n+1} - 2u_j^{n+1} + u_{j-1}^{n+1}\right) + (1-\theta)\left(u_{j+1}^n - 2u_j^n + u_{j-1}^n\right)\right)$$

Defining $\mu = \frac{\varepsilon \Delta t}{(\Delta x)^2}$ and again using $U_j^n = (\lambda)^n e^{ik(j\Delta x)}$ and all it's derived equalities we can rewrite the equation as

$$\lambda u_j^n - u_j^n = \mu\Big(\theta\big(\lambda(e^{ik\Delta x}u_j^n - 2u_j^n + e^{-ik\Delta x}u_j^n)\big) \ldots$$

$$\ldots + (1-\theta)\big(e^{ik\Delta x}u_j^n - 2u_j^n + e^{-ik\Delta x}u_j^n\big)\Big).$$

Again dividing through by u_j^n we have

$$\lambda = 1 - 2\mu\Big(\theta\lambda\big(1-\cos(k\Delta x)\big) + (1-\theta)\big(1-\cos(k\Delta x)\big)\Big),$$

$$\lambda\Big(1 + 2\mu\theta\big(1-\cos(k\Delta x)\big)\Big) = 1 - 2\mu(1-\theta)\big(1-\cos(k\Delta x)\big),$$

$$\lambda = \frac{1 - 4\mu(1-\theta)\left(\sin^2\left(\frac{k\Delta x}{2}\right)\right)}{1 + 4\mu\theta\left(\sin^2\left(\frac{k\Delta x}{2}\right)\right)}.$$

~~Since $\mu >$~~ If we assume $\varepsilon > 0$ then $\mu > 0$, we have also defined θ in the domain $0 \leq \theta \leq 1$. Hence we can never have $\lambda > 1$ and so we only have instability is $\lambda < -1$, therefore

$$1 - 4\mu(1-\theta)\left(\sin^2\left(\frac{k\Delta x}{2}\right)\right) < -\left(1 + 4\mu\theta\left(\sin^2\left(\frac{k\Delta x}{2}\right)\right)\right)$$

is a condition of instability. This simplifies to

$$\mu(1-2\theta)\sin^2\left(\frac{k\Delta x}{2}\right) > \frac{1}{2},$$

and taking the mode $k\Delta x = \pi$ that is most likely to violate the condition we have

$$\mu(1-2\theta) > \frac{1}{2}$$

as the instability condition.

Now, if $\frac{1}{2} \leq \theta \leq 1$ then the stability condition is in the interval $[-\mu, 0]$ and since we assume $\mu > 0$ then this can never be greater than $\frac{1}{2}$, hence the scheme is stable for all μ. In contrast, when we have $0 \leq \theta < \frac{1}{2}$ then the stability condition is in the interval $(0, \mu]$ and therefore the scheme is only stable if and only if

$$\mu \leq \frac{1}{2(1-2\theta)} .$$

We can perform a similar analysis for the advection term by setting $\varepsilon = 0$ and writing out the θ-scheme discretisation for just the advective term,

$$\frac{U_j^{n+1} - U_j^n}{\Delta t} = a\left(\frac{\theta(U_{j+1}^{n+1} - U_j^{n+1})}{\Delta x} + \frac{(1-\theta)(U_{j+1}^n - U_j^n)}{\Delta x} \right).$$

Defining $\mu = \frac{a\Delta t}{\Delta x}$, $a < 0$ and again using ~~the Fourier mode~~ $U_j^n = (\lambda)^n e^{ik(j\Delta x)}$ and it's derived equalities we can rewrite the equation as

$$\lambda U_j^n - U_j^n = \mu\left(\theta(\lambda e^{ik\Delta x} U_j^n - \lambda U_j^n) + (1-\theta)(e^{ik\Delta x} U_j^n - U_j^n)\right),$$

~~[illegible]~~ and again dividing by U_j^n we have

$$\lambda = 1 + \mu\theta\lambda(e^{ik\Delta x} - 1) + \mu(1-\theta)(e^{ik\Delta x} - 1),$$

$$\lambda\left(1 - \mu\theta(e^{ik\Delta x} - 1)\right) = 1 + \mu(1-\theta)(e^{ik\Delta x} - 1),$$

$$\lambda = \frac{1 + \mu(1-\theta)(e^{ik\Delta x} - 1)}{1 - \mu\theta(e^{ik\Delta x} - 1)} .$$

This leads to

$$|\lambda|^2 = \frac{1 - 2\mu(1-\theta)\left[1-\mu(1-\theta)\right]\left[1-\cos(k\Delta x)\right]}{1+2\mu\theta\left[1+\mu\theta\right]\left[1-\cos(k\Delta x)\right]},$$

$$|\lambda|^2 = \frac{1 - 4\mu(1-\theta)\left[1-\mu(1-\theta)\right]\sin^2\left(\frac{k\Delta x}{2}\right)}{1+4\mu\theta\left[1+\mu\theta\right]\sin^2\left(\frac{k\Delta x}{2}\right)}.$$

For stability we must have $|\lambda| \leq 1$ and so we can write an instability condition as

$$\left|1-4\mu(1-\theta)(1-\mu(1-\theta))\sin^2\left(\frac{k\Delta x}{2}\right)\right| > \left|1+4\mu\theta(1+\mu\theta)\sin^2\left(\frac{k\Delta x}{2}\right)\right|.$$

The term on the left-hand side is always positive since $\mu < 0$. Therefore we can split the condition into two cases:

$$1-4\mu(1-\theta)(1-\mu(1-\theta))\sin^2\left(\frac{k\Delta x}{2}\right) > 1+4\mu\theta(1+\mu\theta)\sin^2\left(\frac{k\Delta x}{2}\right),$$

$$(4\mu\theta-4\mu)(1+\mu\theta-\mu)\sin^2\left(\frac{k\Delta x}{2}\right) > 4\mu\theta(1+\mu\theta)\sin^2\left(\frac{k\Delta x}{2}\right),$$

$$-4\mu^2\theta\sin^2\left(\frac{k\Delta x}{2}\right) - 4\mu(1+\mu\theta-\mu)\sin^2\left(\frac{k\Delta x}{2}\right) > 0.$$

Taking the mode most likely to be unstable, i.e. $k\Delta x = \pi$, we have

$$4\mu\left[\mu\theta + 1 - \mu(1-\theta)\right] < 0,$$

$$4\mu(1+2\mu\theta-\mu) < 0,$$

$$\mu(1+\mu(2\theta-1)) < 0.$$

If we have $0 \leq \theta \leq \frac{1}{2}$ then because $\mu < 0$ the scheme is always unstable. If $\frac{1}{2} < \theta \leq 1$ then again since $\mu < 0$ we must have

$$1+\mu(2\theta-1) > 0,$$ which is always satisfied for $\frac{1}{2} < \theta \leq 1$ and again the scheme is unstable.

Taking the other case we have

$$1-4\mu(1-\theta)(1-\mu(1-\theta))\left(\sin^2\left(\frac{k\Delta x}{2}\right)\right) < -\left[1+4\mu\theta(1+\mu\theta)\sin^2\left(\frac{k\Delta x}{2}\right)\right],$$

$$\cancel{2+4\mu^2\sin^2\left(\frac{k\Delta x}{2}\right)+4\mu}$$

$$2+8\mu\theta(1+\mu\theta)\sin^2\left(\frac{k\Delta x}{2}\right)-4\mu^2\theta\sin^2\left(\frac{k\Delta x}{2}\right)-4\mu(1-\mu(1-\theta))\sin^2\left(\frac{k\Delta x}{2}\right) < 0,$$

and again taking the most unstable mode i.e $k\Delta x = \pi$ we have

$$2+8\mu\theta(1+\mu\theta)-4\mu^2\theta-4\mu(1-\mu(1-\theta)) < 0,$$

$$2+8\mu\theta+8\mu^2\theta^2-4\mu^2\theta-4\mu+4\mu^2-4\mu^2\theta < 0,$$

$$8\mu^2\theta^2-8\mu^2\theta+8\mu\theta+4\mu^2-4\mu < -2,$$

$$8\mu\theta(\mu\theta-\mu+1)+4\mu(\mu-1) < -2,$$

$$2\mu\theta(\mu\theta-\mu+1)+\mu(\mu-1) < -\frac{1}{2}.$$