NUMERICAL SOLUTION OF THE STEADY STATE PROPAGATION OF NONLINEAR DIFFUSIVE WAVES

Paula de Oliveira*, Ana Paula Mouro*
* Department of Mathematics, University of Coimbra, 3000 Coimbra, Portugal.

Abstract. The steady state propagation of nonlinear diffusive waves that are a solution of the Boundary Value Problems (BVP) $-\varepsilon u_{xx}+b(x,u)u_x=0$ is studied. When ε approaches a critical value, standard numerical methods used in a uniform mesh with a reasonable stepsize, fail to converge. Using a coordinate transformation we associate to the initial BVP a mesh equation, which leads to an accurate and extremely efficient numerical method. Numerical simulations that show the effectiveness of our approach are included.

1. Introduction. Nonlinear differential equations arise frequently in formulating problems of Science and Engineering. In the framework of the propagation of nonlinear diffusive waves, as in problems involving sound waves in a viscous medium, waves in fluid-filled viscous elastic tubes or magnetohydrodynamic waves in a medium with finite electrical conductivity, the steady state behavior is described by nonlinear Boundary Value Problems of type

$$\textbf{(1.1)}\qquad \begin{cases} -\varepsilon\, u_{xx}+b(x,u)u_x=0, \\ u(0)=A,\, u(1)=B. \end{cases}$$

Problems of type (1.1), depending on a small parameter ε, change abruptly in layers because their solutions approach a discontinuous limit as the small parameter approaches some critical value. These problems are called Singularly Perturbed BVP. When conventional descritization techniques are used to compute the numerical solution of (1.1) in an uniform mesh with a reasonable stepsize, results are very inaccurate when the small parameter is close to such a critical value.

Many authors have studied in these last years numerical methods for singularly perturbed boundary value problems. Essentially two alternative ways of dealing with such problems can be found treated in the literature: (i) the use, on equidistant meshes, of elaborate schemes based on exponential fitting or flux corrected approaches; (ii) the use of conventional methods, on highly non equidistant meshes a priori defined or defined by some adaptive procedure. In the case of meshes a priori defined, or meshes defined by some adaptive procedure the information coming from a previous computation of an approximated solution on a given mesh should be used.

In this paper we suggest a new approach to solve Singularly Perturbed Boundary Value Problems, in highly non equidistant meshes well adapted to the behavior of the solution, while avoiding an “a priori“ knowledge of the qualitative properties of the solution. The idea underlying our procedure is that to solve a problem with a mesh of high density in the layer(s) is equivalent to solving a modified equation with a more regular solution in an equidistant mesh. We consider problems of type (1.1). We then look for a mesh generating function such that the initial BVP is equivalent to a modified problem, which solution is a first order polynomial. This corresponds to the selection of a mesh with a density proportional to the gradient of u. In fact, let $u(x)$ be the solution of (1.1) and let $x=g(\xi)$ be the mesh generating function. We have $u(x)=u(g(\xi))$ that is

$u(x) = (B - A)\xi + A$ for $x = g(\xi)$. We also show that the approximated solution u_i at $x = x_i$ obtained by this procedure that is $u_i = (B - A)\xi_i + A$, where u_i represents an approximation of $u(x)$ at $x_i = g(\xi_i)$, is the solution of (1.1) when solved in a non equidistant mesh and with a certain numerical method.

The paper is organized as follows. In Section 2 we define a mesh density function and construct the mesh equation. In Section 3 a discretization of the mesh equation is studied, and finally in Section 4 some numerical simulations are presented.

2. The mesh equation.

Let us consider the singularly perturbed BVP

(2.1) $$\begin{cases} -\varepsilon u_{xx} + b(x,u)u_x = 0, \\ u(0) = A, u(1) = B, \end{cases} \qquad B \neq A.$$

If we solve (2.1) with a standard numerical method like centered (CFD) or upwind finite-difference approximations results are very inaccurate (Fig.1) when ε is very small.

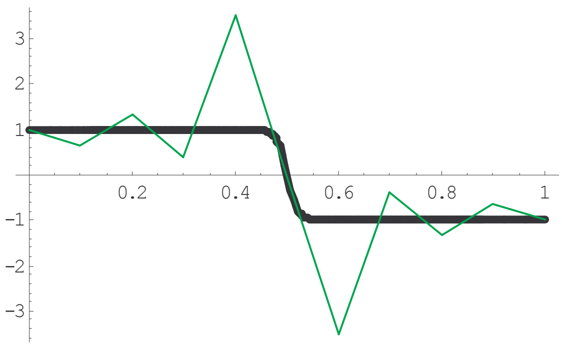


Fig.1 Exact and numerical solutions of (1.1) with $A = 1$, $B = -1$, $\varepsilon = 10^{-2}$, $b(x,u) = u$, N=10, using CFD.

To obtain accurate solutions, presenting no numerical dispersion, centered-finite differences should be used with N=50. In fact it can be proved that to obtain non oscillatory solutions the stepsize must satisfy $h \leq \dfrac{2\varepsilon}{\max|b|}$. In the case of upwind methods as u changes sign at $x = \bar{x}$ with $\bar{x} \in [0,1]$ and $\bar{x}$ is not a priori known, we don´t know where to switch the discretization of u_x from forward to backward or from backward to forward.

In our approach we begin by changing the independent variable by means of a function $g:[0,1] \to [0,1]$ such that $g(0) = 0, g(1) = 1$ and $g_\xi \neq 0$ in $[0,1]$. The modified problem takes then the form

(2.2) $$\begin{cases} -\varepsilon \bar{u}_{\xi\xi} \dfrac{1}{g_\xi^2} + \left[\varepsilon \dfrac{g_{\xi\xi}}{g_\xi^3} + b(\bar{u}(\xi), g(\xi)) \dfrac{1}{g_\xi} \right] \bar{u}_\xi = 0, \\ \bar{u}(0) = A, \quad \bar{u}(1) = B, \end{cases} \qquad B \neq A,$$

where $\bar{u}(\xi) = (u \circ g)(\xi)$. We now select g such that $\bar{u}_{\xi\xi} = 0$ that is such that g is a solution of

(2.3)
$$\begin{cases} \varepsilon\, g_{\xi\xi}(B-A) + \bar{b}(\xi, g) g_{\xi}^{2}(B-A) = 0, \\ g(0) = 0, \quad g(1) = 1, \end{cases}$$

where $\bar{b}(\xi, g) = b((B-A)\xi + A, g(\xi))$.

If $B \neq A$**,** we can rewrite (2.3) in the form

(2.4)
$$\begin{cases} \varepsilon\, g_{\xi\xi} + \bar{b}(\xi, g) g_{\xi}^{2} = 0, \\ g(0) = 0, \quad g(1) = 1. \end{cases}$$

To clarify the meaning of this change of variable let us consider b constant, A=0, B=1. The solution of (2.4) is then

(2.5)
$$g(\xi) = \frac{\varepsilon}{b} \ln \frac{\xi + (e^{\frac{b}{\varepsilon}} - 1)^{-1}}{(e^{\frac{b}{\varepsilon}} - 1)^{-1}}.$$

Fig. 2 shows the plot of g for $\varepsilon = 10^{-2}$, b>0 and b<0. The mesh defined by (2.5), $x_i = g(\frac{i}{N}), i = 1, ..., N-1,$ locates accurately the boundary layers. In Fig.3 we present the solution of (2.1) for the previous meshes.

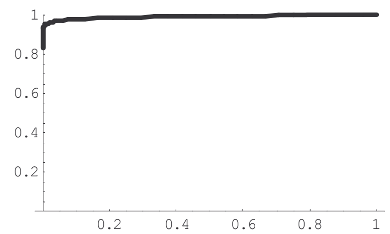


Fig.2a. The plot of g for b=1.

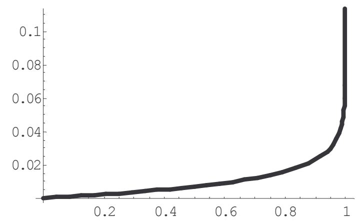


Fig.2b. The plot of g for b=-1.

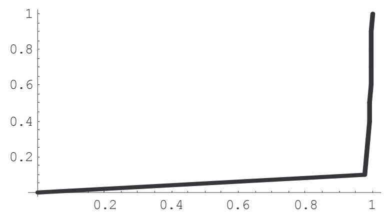


Fig.3a. The solution of (2.1) in the mesh defined by (2.5) with b=1.

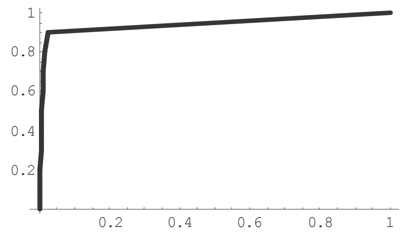


Fig.3b. The solution of (2.1) in the mesh defined by (2.5) with b=-1.

In Table1 we can conclude that for b=1, $g(\xi) = 1 - O(\varepsilon)$, where $\xi_1 = \frac{1}{N}$ and N is the number of nodes considered in the discretization. We observe that all nodes are located in $\left[g(\xi_1), 1 \right]$.

ε	$g(\xi_1)$
10^{-2}	0.9770
10^{-3}	0.9977
10^{-4}	0.9998

Table1. The first node for different values of ε.

An interpretation of the meaning of the mesh function g is also possible in a general case. In fact from (2.4) we have

(2.6) $$-\frac{g_{\xi\xi}}{g_\xi^2}=\frac{\bar{b}(\xi,g)}{\varepsilon}.$$

Let us represent by $d(\xi)$ the function $\frac{1}{g_\xi}$. As $\frac{1}{g_\xi}=\frac{u_x}{B-A}$ function $d(\xi)$ represents a mesh density proportional to the gradient of u. From (2.6) we have

(2.7) $$d_\xi(\xi)=\frac{\bar{b}(\xi,g)}{\varepsilon}.$$

If $\bar{b}(\xi,g)>0, \forall\xi\in[0,1]$ then d is an increasing function of ξ, which means that a boundary layer exists at x=1. On the other hand if $\bar{b}(\xi,g)<0, \forall\xi\in[0,1]$, then d is a decreasing function of ξ and the boundary layer is located at x=0. In the case that $\bar{b}$ changes sign in [0,1] for some ξ_c then d is not a monotone function. Let us consider for instance that $\bar{b}(\xi,g)>0, \forall\xi\in[0,\xi_c[$ and $\bar{b}(\xi,g)<0, \forall\xi\in]\xi_c,1]$. As before we conclude that d is an increasing function in $[0,\xi_c[$ and a decreasing function in $]\xi_c,1]$, which indicates the existence of an internal layer at $\xi=\xi_c$. If we consider $\bar{b}(\xi,g)<0, \forall\xi\in[0,\xi_c[$ and $\bar{b}(\xi,g)>0, \forall\xi\in]\xi_c,1]$, we will expect two boundary layers located at x=0 and x=1.

In any case, the mesh generating function defined by (2.4) provides us a good localization of the boundary layers.

Let us return to equation (2.4). In most cases that equation cannot be solved exactly. In the next section we propose a discretization of (2.4) in uniform meshes.

3. Discretization of the mesh equation.

We begin by considering a discretization of the mesh (2.4) in [0,1] using a uniform stepsize k:

(3.1) $$\varepsilon\frac{g_{i+1}-2g_i+g_{i-1}}{k^2}+\bar{b}_i\frac{(g_{i+1}-g_i)(g_i-g_{i-1})}{k^2}=0, \qquad i=1,\ldots,N-1,$$

where g_i stands for an approximation of $g(\xi_i)$ and $\bar{b}_i=b(\xi_i,g_i)$.

Discretization (3.1) is equivalent to

(3.3) $$\varepsilon(g_{i+1}-2g_i+g_{i-1})+\bar{b}_i(g_{i+1}-g_i)(g_i-g_{i-1})=0.$$

Method (3.3) is locally stable because it can be considered as an upwind type method. We can write (3.3) in the form

(3.4) $$\varepsilon \frac{g_{i+1} - 2g_i + g_{i-1}}{(g_{i+1} - g_i)(g_i - g_{i-1})}(2k)(B-A) + \bar{b}_i(2k)(B-A) = 0.$$

Noting that $\bar{u}_{i+1} - \bar{u}_{i-1} = (B-A)\xi_{i+1} - (B-A)\xi_{i-1}$, we have $\bar{u}_{i+1} - \bar{u}_{i-1} = 2k(B-A)$. On the other hand, as $g_{i+1} - g_{i-1} = h_i + h_{i+1}, g_{i+1} - 2g_i + g_{i-1} = h_{i+1} - h_i, g_{i+1} - g_i = h_i$, we conclude that (3.4) is equivalent to

(3.5) $$\varepsilon \frac{h_{i+1} - h_i}{h_i h_{i+1}}(\bar{u}_{i+1} - \bar{u}_{i-1}) + \bar{b}_i(\bar{u}_{i+1} - \bar{u}_{i-1}) = 0.$$

Finally remarking that $\bar{u}(\xi)$ is a first order polynomial, $\bar{u}_{i+1} - 2\bar{u}_i + \bar{u}_{i-1} = 0$ and we can establish that (3.5) is equivalent to

(3.6) $$\varepsilon\left[-(\bar{u}_{i+1} - 2\bar{u}_i + \bar{u}_{i-1})\frac{h_i + h_{i+1}}{h_i h_{i+1}} + (\bar{u}_{i+1} - \bar{u}_{i-1})\frac{h_{i+1} + h_i}{h_i h_{i+1}}\right] + \bar{b}_i(\bar{u}_{i+1} - \bar{u}_{i-1}) = 0$$

and therefore to

(3.7) $$-\varepsilon \frac{h_i \bar{u}_{i+1} - (h_i + h_{i+1})\bar{u}_i + h_i \bar{u}_{i-1}}{\frac{1}{2}h_i h_{i+1}(h_i + h_{i+1})} + \bar{b}_i \frac{\bar{u}_{i+1} - \bar{u}_{i-1}}{h_i + h_{i+1}} = 0.$$

We have then proved that the approximation $\bar{u}_i$ of $u(x_i)$ defined by $\bar{u}_i = (B-A)\xi_i + A$ is the central finite difference discretization computed in the mesh given by (3.3). Following this approach, instead of solving (2.1) we solve the mesh equation (3.2) obtaining g_i and consequently an approximation $\bar{u}_i$ of $u(g(\xi_i))$, that is $u(x_i)$. As we have mentioned before if we solve directly (2.1) with centered finite difference we must use a very refined mesh.

4. Some properties of the mesh generating function.

In this section we prove that function *g* generates a mesh that satisfies $g \geq 0$.

Define the nonlinear mapping $F: IR^{N-1} \to IR^{N-1}$ by the symmetric of the left--hand side of (3.4). Equation (3.4) is then equivalent to and $g_{i+1} > g_i, \quad i = 1, ..., N-1$.

(4.1) $F(g_1, g_2, ..., g_N) = k$,

where *k* is defined by $k^T = [0 \quad 0 \quad ... \quad \varepsilon]^T$, because $g_0 = 0, g_N = 1$.

To prove that the solution $g_i, i = 1, ..., N$, is positive it is enough to show that *F* is an inverse monotone function ([1]) that is if $Fg \geq 0$ then $g \geq 0$. If the Jacobian matrix of *F*, JF, is inverse monotone, it is obvious, using the Mean Value Theorem for each coordinate function F_i of *F*, that *F* is inverse monotone.

The elements of *JF* are defined by

$$\textbf{(4.2)} \qquad \begin{cases} -\dfrac{\varepsilon}{(g_i - g_{i-1})^2} & j = i-1, \\ \varepsilon \dfrac{(g_{i+1} - g_i)^2 + (g_i - g_{i-1})^2}{(g_{i+1} - g_i)^2 (g_i - g_{i-1})^2} & j = i, \\ -\dfrac{\varepsilon}{(g_{i+1} - g_i)^2} & j = i+1, \\ 0 & \textit{in other cases.} \end{cases}$$

To establish that *JF* is inverse monotone we need the following definitions.

Definition 4.1. The matrix $A=[a_{ij}]$ is called an *M-matrix* if

(4.3a) $a_{ii} > 0$, for all $i \in I$, $a_{ij} \le 0$, for all $i \ne j$.

(4.3b) A is nonsingular and $A^{-1} \ge 0$.

The index $i \in I$ is said to be directly connected with $j \in I$ if $a_{ij} \ne 0$. We said that $i \in I$ is connected with $j \in I$ if there exists a "connection" (chain of direct connections) $i = i_0, i_1, ..., i_k = j$, with $a_{i_{p-1} i_p} \ne 0$ $(1 \le p \le k)$. The index set *I* together with the direct connection form the graph of *A*. In the case that *A* has a symmetrical structure , we have $a_{ij} \ne 0$ if and only if $a_{ji} \ne 0$. In this case *i* is (directly) connected with *j* if and only if *j* is (directly) connected with *i*.

Definition 4.2. A matrix *A* is said to be *irreducible* if every $i \in I$ is connected with every $j \in I$.

Definition 4.3. a) *A* is said to be diagonally dominant if

$$\textbf{(4.4a)} \qquad \sum_{\substack{j \ne i \\ j \in I}} |a_{ij}| < |a_{ii}|, \qquad \text{for all } i \in I.$$

b) *A* is said to be irreducibly diagonally dominant if *A* is irreducible and the inequality (4.4a) holds for at least one index $i \in I$ and

$$\textbf{(4.4b)} \quad \sum_{\substack{j \ne i \\ j \in I}} |a_{ij}| \le |a_{ii}|, \qquad \text{for all } i \in I.$$

In next theorem we prove that *JF* is inverse monotone.

Theorem 4.1. The matrix *JF* is inverse monotone.

Proof. Let $I = \{1, ..., N-1\}$. *JF* is a tridiagonal symmetric matrix. We have $a_{12} \ne 0, a_{N-1,N-2} \ne 0, a_{i,i-1}\, a_{i,i+1} \ne 0, 1 \le i \le N-2$, which means that every $i \in I$ is connected with every $j \in I$. Therefore *JF* is irreducible.

From (4.2), we notice that (4.3a) holds for the matrix *JF*. We have also

(4.5) $$\left|(JF)_{11}\right| > \sum_{\substack{j\neq 1\\ j\in I}} \left|(JF)_{1j}\right|,$$

which means that (4.4b) holds for *i*=1. Therefore *JF* is irreducibly diagonally dominant. On the other hand, the entries of *JF* satisfy (4.3a). Following [3] if a matrix has property (4.3a) and is irreducibly diagonally dominant then it is an M-matrix.

We conclude then that *JF* is an M-matrix and then *JF* is invertible and $(JF)^{-1} \geq 0$.

If *JF* is an M-matrix, then *JF* is invertible and $(JF)^{-1} \geq 0$.

■

To guarantee that g generates a mesh we must prove firstly that $g_i > 0, \quad i = 1, ..., N-1,$ and secondly that $g_i < g_{i+1}, \quad i = 0, ..., N-1.$ In fact

Theorem 4.2. The mesh function *g* is positive.

Proof. Since $Fg = \begin{bmatrix} 0 & 0 & \cdots & 0 & \varepsilon \end{bmatrix}^T \geq 0$ and *JF* is inverse monotone, *F* is an inverse monotone operator and $Fg \geq 0$ implies $g \geq 0$.

■

Using the previous proposition we easily establish that

Theorem 4.3. The mesh function *g* satisfies $g_i < g_{i+1}, \quad i = 0, ..., N-1.$

Proof. Let $v_i = g_i - g_{i-1}$. From (3.3), we have

(4.6) $$\varepsilon(v_{i+1} - v_i) + \bar{b}_i v_i v_{i+1} = 0,$$

or, equivalently,

(4.7) $$\frac{v_{i+1} - v_i}{v_i v_{i+1}} = -\frac{\bar{b}_i}{\varepsilon}.$$

If $\bar{b}_i > 0$, since $\varepsilon > 0$, we have $\frac{v_{i+1} - v_i}{v_i v_{i+1}} < 0$, and consequently $v_{i+1} - v_i < 0$ and $v_i v_{i+1} > 0$ or $v_{i+1} - v_i > 0$ and $v_i v_{i+1} < 0$. As we already proved that $g_1 > 0$ we easily conclude after some computations that $g_i < g_{i+1}$.

If $\bar{b}_i < 0$ then the same result holds.

In both cases we have $v_i > 0$ for all $i \in I$, which means that $0 \equiv g_0 < g_1 < ... < g_{N-1} < g_N \equiv 1$. Therefore *g* is an increasing function.

■

5. Numerical results.

Example 1. Let us consider

$$\text{(5.1)} \quad \begin{cases} -\varepsilon u_{xx} + (u-a)u_x = 0, \\ u(0) = 0, \;\; u(1) = 1, \end{cases}$$

with $0 \le a < 1$ and $\varepsilon = 10^{-2}$. For a=0, we have the Burgers equation. Fig.4 shows the behavior of the solution for different values of a, using the mesh equation (3.3). We can observe, for some values of a, two boundary layers located at x=0 and x=1. If we take a=0, then we have only one boundary layer located at x=1.

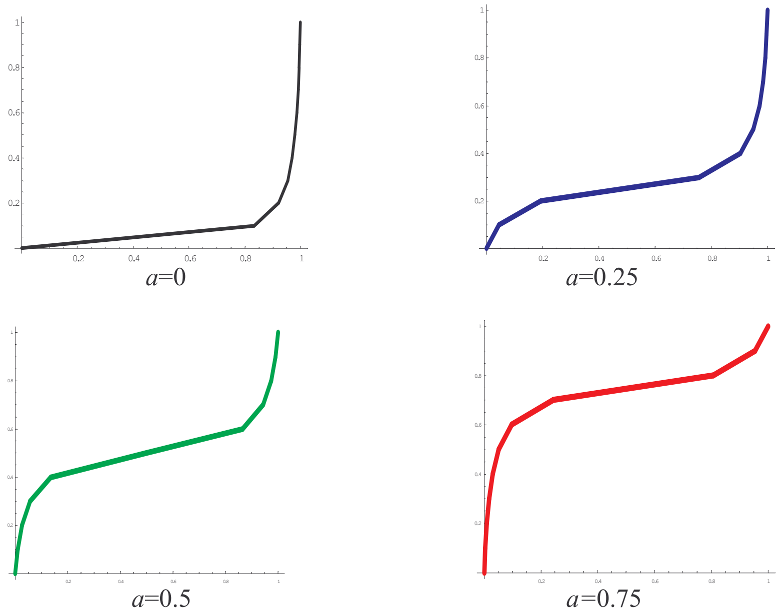


Fig.4 The numerical solution of (5.1) for a=0, 0.25, 0.5, 0.75,using the mesh obtained by (3.4).

In Table2 we indicate the mesh obtained by solving (5.1) for a=0.5, $\varepsilon = 10^{-2}$, using (3.4).

	Mesh
a=0.5	0.0097, 0.0257, 0.0562, 0.1347, 0.5000, 0.8653, 0.9438, 0.9743, 0.9903

Table2. The mesh obtained to the problem (5.1) using (3.4).

Example 2. Let us consider

$$\textbf{(5.2)} \quad \begin{cases} -\varepsilon u_{xx} + (0.5-u)u_x = 0, \\ u(0)=0, \ \ u(1)=1. \end{cases}$$

Fig.5 shows the behavior of the solution of (5.1) using the mesh obtained using (3.4). In this case, we can observe an internal layer.

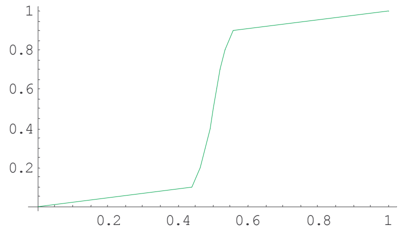


Fig.5 The numerical solution of (5.2), using the mesh obtained by (3.4).

In Table3 we indicate the mesh obtained by solving (5.2) for $\varepsilon = 10^{-2}$, using (3.4).

Mesh
0.4420, 0.4656, 0.4795 0.4903, 0.5001, 0.5099, 0.5207, 0.5345, 0.5582

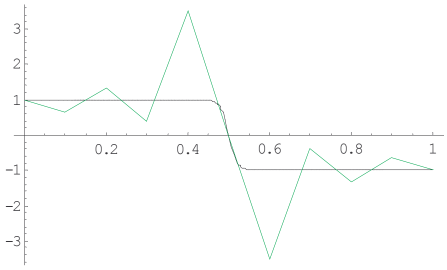
Table2. The mesh obtained for the problem (5.1) using (3.4).

References

[1] H.-G. Roos, M. Stynes, L. Tobiska, Numerical methods for singularly perturbed differential equations: convection-diffusion and flow problems, Springer-Verlag, 1996.

[2] Lokenath Debnath, Nonlinear partial differential equations for scientists and engineers, Birkhäuser, 1997.

[3] W. Hackbusch, Elliptic differential theory and numerical treatment, Springer-Verlag, 1992.


3
2
1
-1
-2
-3
0.2
0.4
0.6
0.8
1