

Optimal Upstream Collocation Solution of a Convection-Diffusion Equation

Stephen H. Brill

Boise State University
Department of Mathematics
Boise, ID 83725-1555
<http://math.boisestate.edu/~brill>

Abstract

We give herein analytical formulas for the solution of the Hermite collocation discretization of the unforced steady-state convection-diffusion equation in one spatial dimension and with constant coefficients, defined on a uniform mesh with Dirichlet boundary conditions. The accuracy of the method is enhanced by employing “upstream weighting” of the convective term in an optimal way, avoiding both the “smearing” effect of numerical diffusion and unwanted oscillations, particularly for large Péclet numbers. Computational examples illustrate the efficacy of using optimal upstream weighting.

1 Introduction

It is well known that the numerical solution of convection-diffusion differential equations (DEs) is a difficult task when convection is the dominant process. Numerical techniques often give rise to spurious oscillations that are not present in the continuous (i.e., not numerical/discrete) solution of the DE. To ameliorate these physically unmeaningful (and therefore undesirable) oscillations, the technique of upstream weighting is often used ([1], [3]). While upstreaming can eliminate the oscillations, it is often at the expense of “smearing” the sharp solution profile of the continuous solution of the DE.

In this work, we study the convection-diffusion equation

$$-D \frac{d^2 u}{dx^2} + v \frac{du}{dx} = 0 \quad (1)$$

with Dirichlet boundary conditions, defined on the interval $[0, 1]$. The convection coefficient v and diffusion coefficient D are both positive constants. For the purposes of numerical solution, we subdivide the domain $[0, 1]$ into m equal subintervals and thus seek to solve (1) at the nodes $x_j = jh$, $j = 0, 1, 2, \dots, m$, where $h = 1/m$.

The following example, depicted in Figure 1, is discussed in detail in [3] and illustrates the issues that this paper tackles. Suppose along with (1) we have the boundary conditions

$$\begin{aligned} u(0) &= 0 \\ u(1) &= 1. \end{aligned} \quad (2)$$

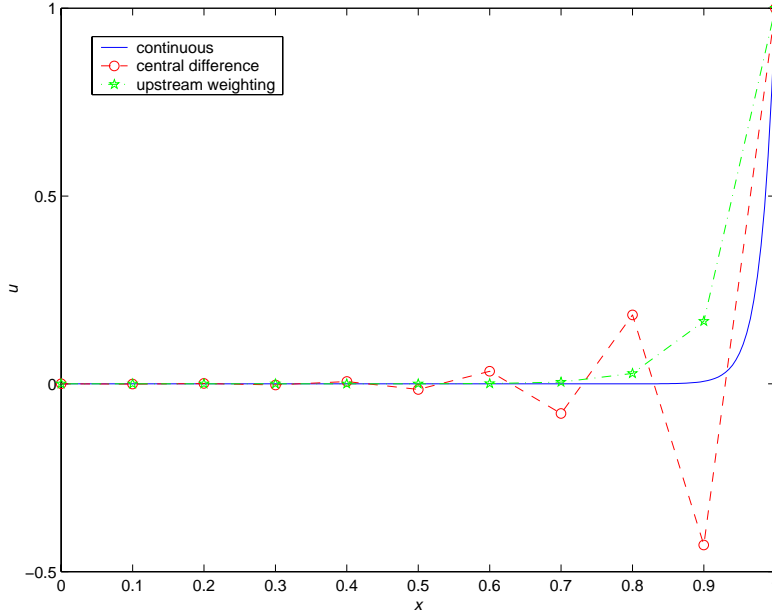


Figure 1: Continuous, central difference, and upstream weighting solutions of (1),(2) with $\beta = 5$ and $m = 10$.

If we discretize (1) via standard central differences, then we obtain the solution

$$u_j^C = \frac{\lambda_C^j - 1}{\lambda_C^m - 1},$$

$j = 0, 1, 2, \dots, m$, where u_j^C is an approximation to the exact solution of the continuous problem, namely

$$u(x_j) = \frac{e^{\beta j} - 1}{e^{\beta m} - 1}. \quad (3)$$

The lumped parameter $\beta = \frac{hv}{D}$ is known as the Péclet number and

$$\lambda_C = \frac{2 + \beta}{2 - \beta}.$$

It is clear that if $\beta > 2$, then $\lambda_C < -1$. So in this case u_j^C and u_{j+1}^C have opposite signs; i.e. the central difference solution oscillates, which is qualitatively very different from the monotone exact solution (3) (see Figure 1).

We may eliminate these oscillations via upstream weighting. For the case of finite difference discretization, this is most easily accomplished by replacing the central difference approximation to $\frac{du}{dx}$ in (1) with a one-sided backward difference approximation. The resulting solution is now

$$u_j^U = \frac{\mu^j - 1}{\mu^m - 1},$$

where $\mu = 1 + \beta$. While this last solution is monotone, we have lost the sharp solution profile that is present in (3) (see Figure 1). This example clearly illustrates the limitations of the finite difference discretization to solve (1). We are thus motivated to investigate other methods of discretization.

This paper is organized as follows. We begin by providing the analytical solution of the matrix equation that arises from the Hermite collocation discretization of the DE (1), when upstream weighting is utilized. Subsequently, we provide an analysis which compares the discrete collocation solution to the continuous solution. In particular, we will discuss how to select the upstream parameter ζ as a function of the Péclet number β in such a way as to eliminate spurious oscillations and minimize the difference between the continuous and discrete solutions. We then provide several computational examples which illustrate the theory. A short section summarizing our results concludes the paper.

2 Analytical Solution of Upstream Collocation

In this work, we discretize (1) by using Hermite collocation combined with upstream weighting of the convective term of (1), as was introduced by Allen [1]. Algebraic manipulation of the system of linear algebraic equations that arises from the discretization leads to an equivalent system with the repeated computational molecule

$$M \begin{bmatrix} q_j \\ r_j \\ q_{j+1} \\ r_{j+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad (4)$$

$j = 0, 1, 2, \dots, m - 1$. Here q_j is an approximation for $u(x_j)$ and r_j is an approximation for $\frac{du}{dx}(x_j)$ for $j = 0, 1, 2, \dots, m$. Note that the matrix equation represented by (4) is a system of $2m$ equations in $2(m + 1)$ unknowns. M is given by

$$M = \begin{bmatrix} 0 & \lambda_{\text{num}} & 0 & -\lambda_{\text{den}} \\ -\beta m(1 - 6\zeta^2) & 1 + \beta\zeta + 3\beta\zeta^2 & \beta m(1 - 6\zeta^2) & -1 - \beta\zeta + 3\beta\zeta^2 \end{bmatrix}.$$

The parameter ζ controls how much upstream weighting is applied. It can vary from zero (in which case there is no upstreaming) to a maximum value of $\frac{1}{2} - \frac{1}{\sqrt{12}}$. Finally, λ_{num} and λ_{den} are, respectively, the numerator and denominator of (7).

We now give the solution of the matrix equation (4). The proof of this result is completely straightforward, though computationally tedious.

Theorem 1. *The general solution of (4) is*

$$q_j = c_1 + c_2 \lambda^j \quad (5)$$

$$r_j = \rho c_2 \lambda^j, \quad (6)$$

where c_1 and c_2 are constants determined by boundary conditions, $\rho = \frac{2\beta m(1+\beta\zeta)}{\beta^2\zeta^2+4\beta\zeta+2}$, and

$$\lambda = \frac{\beta^2 + 6\beta + 12 + 6\beta\zeta(4 + \beta + \beta\zeta)}{\beta^2 - 6\beta + 12 + 6\beta\zeta(4 - \beta + \beta\zeta)}. \quad (7)$$

We note that this result reduces, for the case $\zeta = 0$, to the analogous formulas in [2], in which upstreaming was not considered.

Since we are given the Dirichlet boundary conditions $u(0) = q_0 = b_0$ and $u(1) = q_m = b_1$ we conclude from (5) with $j = 0, m$, that $c_2 = \frac{b_1 - b_0}{\lambda^m - 1}$ and $c_1 = b_0 - \frac{b_1 - b_0}{\lambda^m - 1}$. Thus

$$q_j = b_0 + (b_1 - b_0) \frac{\lambda^j - 1}{\lambda^m - 1}. \quad (8)$$

The solution of the corresponding continuous problem is

$$u(x_j) = b_0 + (b_1 - b_0) \frac{e^{\beta j} - 1}{e^{\beta m} - 1}. \quad (9)$$

3 Oscillations

That oscillations are undesirable is illustrated in the following example. Suppose the solution u of (1) represents the concentration of a contaminant in water, with $u = 0$ representing pristine water and $u = 1$ representing pure contaminant. Then oscillations (like those in Figure 1) may produce values of u less than 0 and/or greater than 1, which are physically absurd.

With respect to (5) and (6), it is clear that our collocation solution will oscillate if and only if $\lambda < 0$ in (7). Keeping in mind that $\beta \in (0, \infty)$ and $\zeta \in [0, \frac{1}{2} - \frac{1}{\sqrt{12}}]$, a bit of algebra shows that:

Theorem 2. *The upstream collocation solution (5) (6) of (1) will exhibit oscillations if and only if $\beta > 6 + 4\sqrt{3}$ and*

$$\frac{1}{2} - \frac{2}{\beta} - \frac{\sqrt{\beta^2 - 12\beta + 24}}{\sqrt{12}\beta} < \zeta \leq \frac{1}{2} - \frac{1}{\sqrt{12}}. \quad (10)$$

This situation is depicted in the small shaded region in the upper part of Figure 2. Thus, given β , we may avoid oscillations by choosing ζ so that the point (β, ζ) lies outside of this shaded region.

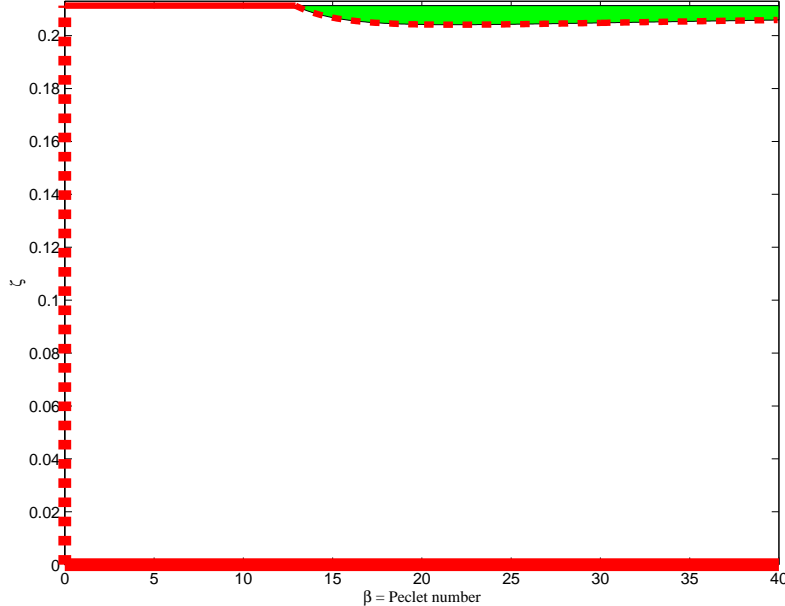


Figure 2: Shaded region indicates where $\lambda < 0$. Outlined region indicates where $\lambda > 0$.

4 Optimal Upstream Weighting

Since we are armed with (8) and (9), we may determine, for a given Péclet number β , the corresponding value of ζ that will minimize the error in using the discrete collocation solution of (1) as an approximation of the corresponding continuous solution. This is done by forcing the derivative of the error with respect to ζ to vanish. For certain values of β , this technique provides the corresponding value of ζ that minimizes the error. For other values of β , however, the error is minimized at the endpoints of the domain of ζ , which is $\left[0, \frac{1}{2} - \frac{1}{\sqrt{12}}\right]$. However, if we also want to eliminate any chance of oscillations, we must ensure that $\zeta < \frac{1}{2} - \frac{2}{\beta} - \frac{\sqrt{\beta^2 - 12\beta + 24}}{\sqrt{12}\beta}$. This is a concern only if $\beta > 6 + 4\sqrt{3}$. In this case, we select ζ to be

$$\zeta = \frac{1}{2} - \frac{2}{\beta} - \frac{\sqrt{\beta^2 - 12\beta + 24}}{\sqrt{12}\beta} - \epsilon, \quad (11)$$

where ϵ is a small positive number large enough to ensure that the computed value of the right side of (11) is indeed less than $\frac{1}{2} - \frac{2}{\beta} - \frac{\sqrt{\beta^2 - 12\beta + 24}}{\sqrt{12}\beta}$.

We thus have an algorithm for choosing ζ optimally:

Theorem 3. *Let ϵ be a small positive number. Assuming that oscillations in the collocation solution are unacceptable, the value of ζ (as a function of β) that minimizes the maximum difference between the discrete collocation solution of (1) and the corresponding continuous solution (both with given Dirichlet boundary conditions) is given in Table 1 (and depicted in Figure 3).*

Table 1: Optimal ζ as a function of β .

β interval	approx β interval	optimal ζ
$(0, 2\sqrt{3}]$	$(0, 3.46410]$	0
$[2\sqrt{3}, \sqrt{3} + 2^{-1/2}(3^{3/4} + 3^{5/4})]$	$[3.46410, 6.13572]$	$\frac{\sqrt{6\beta^2 - 36} - 6}{6\beta}$
$[\sqrt{3} + 2^{-1/2}(3^{3/4} + 3^{5/4}), 6 + 4\sqrt{3}]$	$[6.13572, 12.9282]$	$\frac{1}{2} - \frac{1}{\sqrt{12}}$
$[6 + 4\sqrt{3}, \infty)$	$[12.9282, \infty)$	$\frac{1}{2} - \frac{2}{\beta} - \frac{\sqrt{\beta^2 - 12\beta + 24}}{\sqrt{12}\beta} - \epsilon$

5 Numerical Experiments

In this section we give the results of numerical experiments that illustrate the theoretical conclusions given in Theorem 3, Table 1, and Figure 3. In particular, we examine how well optimal upstream collocation removes the “smearing” effect initially illustrated in Figure 1. (We noted in [2] that, even for fairly modest Péclet numbers, there was significant smearing in the collocation solution of (1) without upstreaming.) The results are depicted in Figure 4, where for Péclet numbers 2, 5, 10, and 40, we plotted the continuous, optimal upstream collocation, and non-upstream collocation solutions for $m = 10$ and the boundary conditions $u(0) = 1$, $u(1) = 0$. When applicable, we used $\epsilon = 10^{-6}$ (see Theorem 3). For very modest Péclet numbers (e.g., $\beta = 2$), we see that all three solutions are visually indistinguishable (in fact, since $\zeta = 0$ is optimal for this case, the two collocation solutions are identical). For $\beta = 5$, we see some smearing of both collocation solutions, but about only half as much in the optimal case as compared to the $\zeta = 0$ case. For the Péclet numbers 10 and 40, we see significant smearing of the collocation solution without upstreaming, but the continuous and optimal upstream collocation solutions are visually indistinguishable.

In Figure 5 we examine how profoundly optimal upstream collocation outperforms collocation without upstreaming. We define improvement as

$$\text{improvement} = \frac{\text{maximum error without upstreaming}}{\text{maximum error with optimal upstreaming}}$$

and compute this ratio for β between 0.5 and 10,000. (ϵ in Theorem 3 is 10^{-6} here). Although the graph in Figure 5 is for $m = 10$, this curve is visually indistinguishable when compared to those curves for $m = 100$ or $m = 1000$. We see that we obtain significant improvement for all values of the Péclet number other than those for which the optimal value of ζ is 0. In particular, for $\beta \geq 14$, we find improvement to be on the order of half a million.

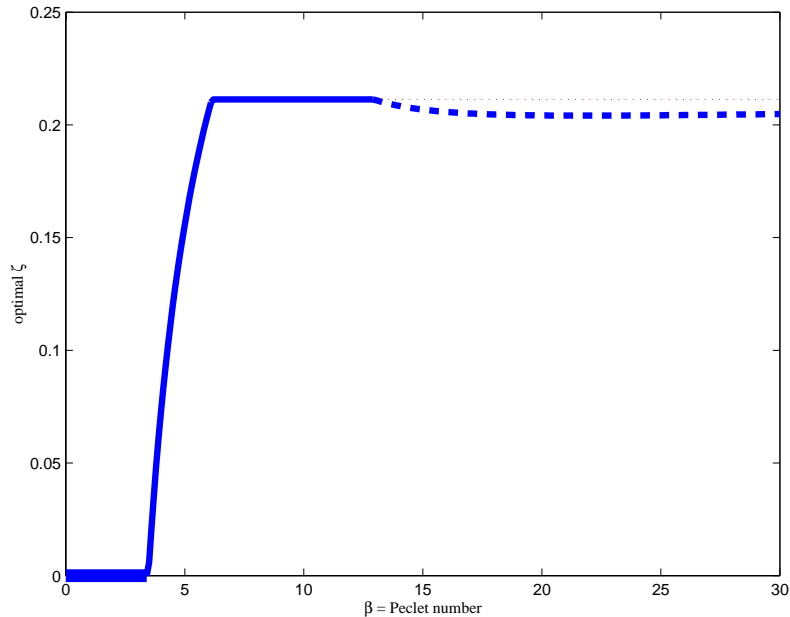


Figure 3: Optimal ζ as a function of Péclet number β

6 Summary and Conclusions

In this paper, we give formulas for the analytical solution of the Hermite collocation discretization of the one-dimensional constant-coefficient unforced convection-diffusion equation, defined on a uniform mesh with Dirichlet boundary conditions. Upstream weighting is employed in the evaluation of the derivative of the convective term. The upstream parameter ζ may be chosen in an optimal manner to both eliminate physically absurd oscillations and to capture the sharp solution profile that exists in the exact solution of the corresponding continuous problem. Numerical experiments conform to and illustrate the theory derived herein.

References

- [1] M. B. Allen, *How Upstream Collocation Works*, Int. J. Num. Meth. Eng., **19** (1983), 1753–1763.
- [2] S. H. Brill, *Analytical Solution of Hermite Collocation Discretization of the Steady-State Convection-Diffusion Equation*, International Journal of Differential Equations and Applications, to appear.
- [3] K. W. Morton, *Numerical Solution of Convection-Diffusion Problems*, Chapman & Hall, London, 1996.

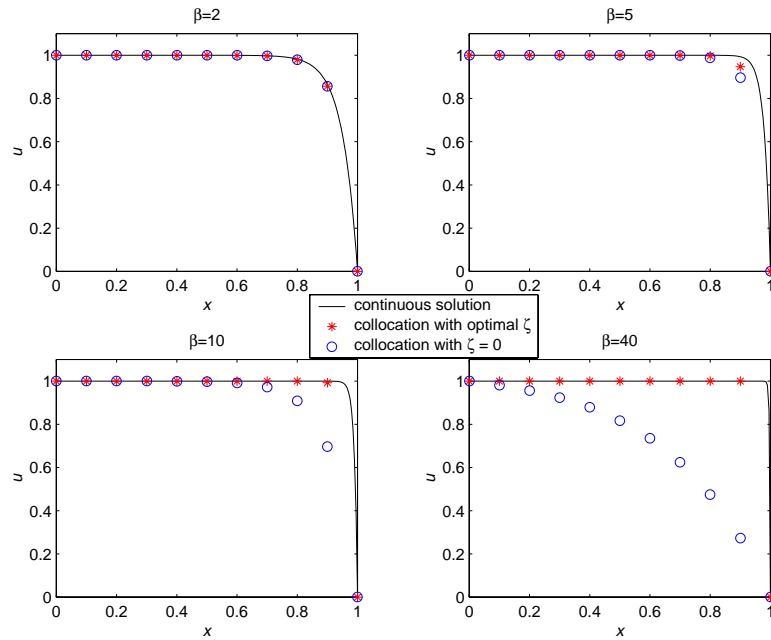


Figure 4: Comparison of continuous solution, optimal upstream collocation solution, and collocation solution with no upstreaming ($\zeta = 0$), for $\beta = 2, 5, 10, 40$.

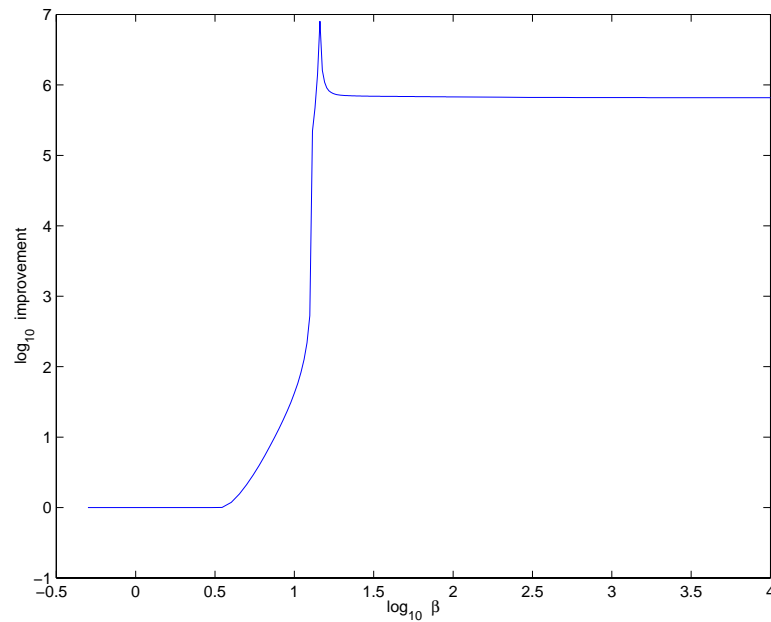


Figure 5: Improvement obtained by using optimal upstream collocation compared to using collocation with no upstreaming. β here varies from 0.5 to 10,000.