# The Analytical Solution of the Hermite Collocation Discretization of the One-Dimensional Steady-State Convection-Diffusion Equation: A Survey of Recent Results

**Stephen H. Brill**
Boise State University
Department of Mathematics
Boise, Idaho 83725-1555
email: brill@math.boisestate.edu

**Abstract**

We give herein analytical formulas for the Hermite collocation solution of the steady-state convection-diffusion equation with constant coefficients and Dirichlet boundary conditions defined on a uniform mesh in one spatial dimension. Analysis is provided which compares the discrete collocation solution to the corresponding exact solution. Extremely accurate collocation solutions are easily attainable in a variety of settings by utilizing the author's "optimal upstream weighting" technique.

# 1 Introduction and background

Convection-diffusion (C-D) differential equations (DEs) are used extensively to study physical processes in the sciences and engineering, including the model of subsurface contaminant transport. The numerical solution of such equations can, however, be plagued by spurious (and physically unmeaningful) oscillations, particularly when convection is the dominant process. To ameliorate the effect of these oscillations, the technique of "upstream weighting" is often applied to the convective term. This introduces artificial dispersion, often at the expense of "smearing" the sharp solution profile that characterizes the convection-dominated problem. The goal, therefore, is to obtain highly accurate solutions that suffer from neither oscillations nor "smearing."

In this paper, we provide formulas for the Hermite collocation solution of the steady-state convection-diffusion equation

$$-D\frac{d^2u}{dx^2} + v\frac{du}{dx} = S(x), \tag{1}$$

with boundary conditions

$$\begin{aligned} u(0) &= u_L \\ u(1) &= u_R, \end{aligned} \tag{2}$$

where the diffusion coefficient $D$ and velocity coefficient $v$ are both positive constants. These formulas are analyzed so that judicious choices can be made concerning the value to be assigned to free coefficients so that highly accurate approximations to the exact solution of our boundary value problem (BVP) can be obtained.

The motivation for choosing the Hermite collocation discretization is that collocation provides great flexibility with respect to where to evaluate the various terms in the discretized version of (1).

Although collocation has been widely studied, there had been no work done in which analytical formulas for collocation discretizations of differential equations were derived, until we turned our attention to this particular problem. In addition to our efforts on the DE (1) (which will be discussed in more detail immediately below), we have studied the one-dimensional self-adjoint constant-coefficient DE [2].

Our initial effort [3] on studying the DE (1) concerned the case where the source/sink term $S(x) \equiv 0$. In this paper, no upstream weighting was considered. In this case, we proved that oscillations were eliminated but that significant "smearing" occurred in the convection dominated case. To eliminate the unacceptable smearing, we next studied the effect of including upstream weighting of the convective term [4], which is governed by a parameter $\zeta$. An algorithm for how to choose $\zeta$ in an optimal fashion was derived. Utilizing this algorithm provides extremely accurate collocation solutions, especially when convection dominates.

We then turned our attention to the case where the source/sink term $S(x)$ is non-zero [5], and derived formulas for the collocation solution for this case. The issue of where to evaluate the source/sink term $S(x)$ arises naturally. For the case where this term is a linear function, we prove that $S(x)$ should be evaluated at precisely the same upstream locations as the convective term in (1) in order to obtain extremely accurate collocation solutions.

The rest of this paper is organized to have one section correspond to each of the three versions of (1) described above. Description of the Hermite collocation technique of discretization is included when appropriate. We conclude this work with a short section summarizing our results.

# 2 No forcing, no upstream weighting

In this section, we assume that the source sink term $S(x)$ is identically zero. Detailed derivations of the results reported in this section may be found in [3]. Upstream weighting will be introduced in the next section.

The differential equation (1) is defined on the interval $\mathcal{I} = [0,1]$. The Dirichlet boundary conditions (2) are included. We partition the interval $\mathcal{I}$ into $m$ uniform subintervals, each of length $h$, by $0 = x_0, x_1, x_2, \ldots, x_m = 1$. Note $x_j = jh$ and $mh = 1$.

The discretization proceeds by introducing a Hermite piecewise cubic interpolating polynomial

$$\widehat{u}\,(x) = \sum_{j=0}^{m} \left[ u_j f_j\,(x) + u_j' g_j\,(x) \right]. \tag{3}$$

into the differential equation (1), obtaining

$$-D\frac{d^2\widehat{u}}{dx^2}+v\frac{d\widehat{u}}{dx}=E(x), \tag{4}$$

where $E(x)$ is an error function.

The Hermite basis functions, defined for $\eta\in\left[-\frac{1}{2},\frac{1}{2}\right]$, are

$$f_j\left(x\right)=\begin{cases} f_L(\eta)=\frac{1}{2}\left(1+2\eta\right)^2\left(1-\eta\right), & x_{j-1}\le x=x_j+\left(\eta-\frac{1}{2}\right)h\le x_j \\ f_R(\eta)=\frac{1}{2}\left(1-2\eta\right)^2\left(1+\eta\right), & x_j\le x=x_j+\left(\eta+\frac{1}{2}\right)h\le x_{j+1} \\ 0, & \text{otherwise}\end{cases} \tag{5}$$

and

$$g_j\left(x\right)=\begin{cases} g_L(\eta)=\frac{h}{8}\left(2\eta+1\right)^2\left(2\eta-1\right), & x_{j-1}\le x=x_j+\left(\eta-\frac{1}{2}\right)h\le x_j \\ g_R(\eta)=\frac{h}{8}\left(2\eta-1\right)^2\left(2\eta+1\right), & x_j\le x=x_j+\left(\eta+\frac{1}{2}\right)h\le x_{j+1} \\ 0, & \text{otherwise.}\end{cases} \tag{6}$$

Note that $\widehat{u}$ in (3) interpolates the values $u_j=u(x_j)$ and $u'_j=\frac{du}{dx}(x_j)$, $j=0,1,\ldots,m$, because $f_j\left(x_k\right)=\delta_{jk}$, $\frac{df_j}{dx}\left(x_k\right)=0$, $g_j\left(x_k\right)=0$, and $\frac{dg_j}{dx}\left(x_k\right)=\delta_{jk}$. Here $\delta_{jk}$ is the Kronecker symbol.

It is clear that (4) has $2\left(m+1\right)$ coefficients, namely $u_j$ and $u'_j$, $j=0,1,2,\ldots m$. However, the imposition of boundary conditions reduces this number to $2m$. To generate the $2m$ equations necessary to find these undetermined coefficients, we enforce that the error function $E(x)$ in (4) is identically zero at two distinct "collocation points" in the interior of each of the $m$ subintervals.

Given certain smoothness conditions, the optimal (in terms of minimizing discretization error) location of the collocation points within each subinterval corresponds to the points of Gaussian quadrature [6]. In this section, we will use these optimal collocation points, which correspond to choosing the collocation points as $\eta=\pm\frac{1}{\sqrt{12}}$ (see (5) and (6)) in each subinterval $\left[-\frac{1}{2},\frac{1}{2}\right]$ (given in local $\eta$ coordinates). (In the following sections, we will no longer collocate all terms of (4) at the Gauss points.)

It is straightforward to see that choosing the collocation points in this manner leads to a matrix equation with the repeated computational molecule

$$\begin{bmatrix} \frac{2\sqrt{3}\,D}{h^2}-\frac{v}{h} & \frac{(1+\sqrt{3})D}{h}+\frac{v}{2\sqrt{3}} & \frac{-2\sqrt{3}\,D}{h^2}+\frac{v}{h} & \frac{(-1+\sqrt{3})D}{h}-\frac{v}{2\sqrt{3}} \\ \frac{-2\sqrt{3}\,D}{h^2}-\frac{v}{h} & \frac{(1-\sqrt{3})D}{h}-\frac{v}{2\sqrt{3}} & \frac{2\sqrt{3}\,D}{h^2}+\frac{v}{h} & \frac{(-1-\sqrt{3})D}{h}+\frac{v}{2\sqrt{3}} \end{bmatrix}\begin{bmatrix} q_j \\ r_j \\ q_{j+1} \\ r_{j+1}\end{bmatrix}=\begin{bmatrix}0\\0\end{bmatrix} \tag{7}$$

for $j=0,1,2,\ldots,m-1$. Here $q_j=u_j$ and $r_j=u'_j$, $j=0,1,2,\ldots,m$. Note that the matrix equation represented by (7) is a system of $2m$ equations in $2(m+1)$ unknowns.

It is equally straightforward, although computationally tedious, to show that the solution of (7) is

$$q_j = \frac{u_L(\lambda^m - \lambda^j) + u_R(\lambda^j - 1)}{\lambda^m - 1} \tag{8}$$

and

$$r_j = \frac{\beta m \lambda^j}{\lambda^m - 1}(u_R - u_L), \tag{9}$$

where

$$\lambda = \frac{\beta^2 + 6\beta + 12}{\beta^2 - 6\beta + 12}. \tag{10}$$

It is instructive to compare the collocation solution (8) and (9) of our BVP with its exact solution and its derivative, both evaluated at $x_j$:

$$u(x_j) = \frac{u_L(e^{\beta m} - e^{\beta j}) + u_R(e^{\beta j} - 1)}{e^{\beta m} - 1} \tag{11}$$

and

$$u'(x_j) = \frac{\beta m e^{\beta j}}{e^{\beta m} - 1}(u_R - u_L) \tag{12}$$

We see that (8) and (9) are, respectively, extremely similar to (11) and (12). The only difference is that the role of $e^{\beta}$ in (11) and (12) is assumed by $\lambda$ in (8) and (9). For an analysis comparing $\lambda$ and $e^{\beta}$, the interested reader is referred to [3]. The most interesting result that that is proven in [3] is that the solution collocation solution is oscillation-free. However, for Péclet numbers of even modest size, the collocation solution suffers from significant smearing, as illustrated in Figure 1.

# 3 No forcing, optimal upstream weighting

In this section we report on results published in [4]. We discuss upstream weighting (originally introduced in [1] for the context of collocation) and its optimal implementation. The advantage of optimal upstream weighting is that we may obtain extremely accurate collocation solutions while still avoiding unwanted oscillations.

When implementing upstreaming, we still enforce $E(x) = 0$ (see (4)) for each of our $2m$ equations and we still evaluate $\frac{d^2\widehat{u}}{dx^2}$ in (4) at the Gaussian points $\eta = \pm\frac{1}{\sqrt{12}}$. However, we evaluate $\frac{d\widehat{u}}{dx}$ at the points $\eta = \pm\frac{1}{\sqrt{12}} - \zeta$, where $\zeta \geq 0$ controls how much upstreaming occurs. Because the support of each basis function $f_j$ or $g_j$ (see (5) and (6)) is the interval $[-\frac{1}{2}, \frac{1}{2}]$, it is clear that $\zeta$ must lie in the interval $[0, \frac{1}{2} - \frac{1}{\sqrt{12}}]$.

It is straightforward to see that collocating in this manner leads to a matrix equation with the repeated computational molecule

$$\begin{bmatrix} M_{11} & M_{12} & -M_{11} & M_{14} \\ M_{21} & M_{22} & -M_{21} & M_{24} \end{bmatrix} \begin{bmatrix} q_j \\ r_j \\ q_{j+1} \\ r_{j+1} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \tag{13}$$

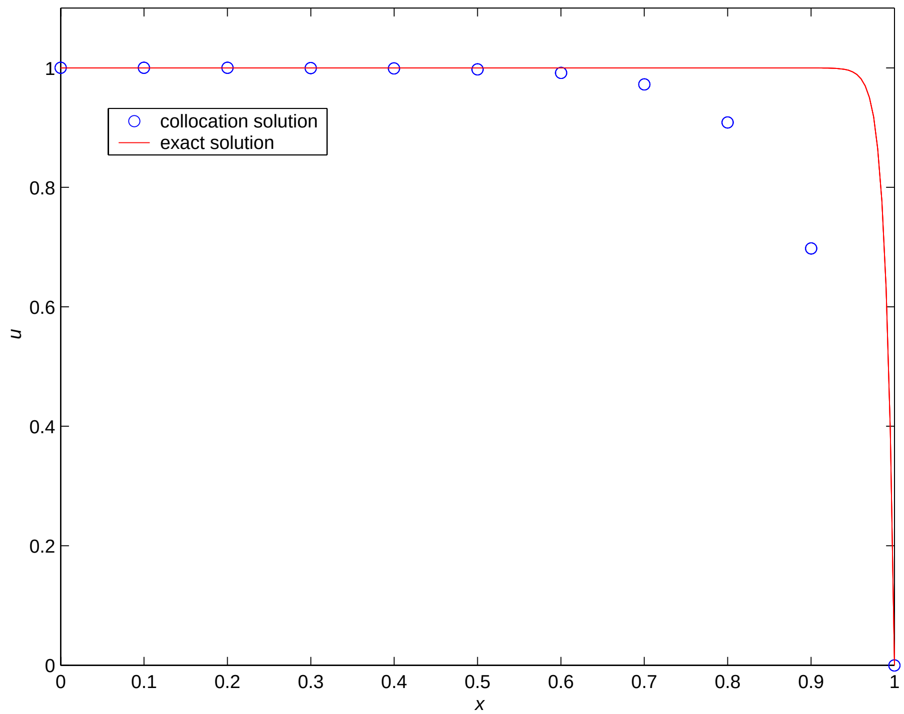


Figure 1: Collocation and exact solutions for $m = 10$, $\beta = 10$, $u_L = 1$, $u_R = 0$, with no forcing or upstream weighting.

The entries of the matrix are

$$
\begin{aligned}
M_{11} &= \frac{2\sqrt{3}D}{h^2} + \frac{v}{h}\left(6\zeta^2 + 2\sqrt{3}\zeta - 1\right) \\
M_{21} &= -\frac{2\sqrt{3}D}{h^2} + \frac{v}{h}\left(6\zeta^2 - 2\sqrt{3}\zeta - 1\right) \\
M_{12} &= \frac{D}{h}(1+\sqrt{3}) + v\left(\frac{\sqrt{3}}{6} + \zeta + \sqrt{3}\zeta + 3\zeta^2\right) \\
M_{22} &= \frac{D}{h}(1-\sqrt{3}) + v\left(-\frac{\sqrt{3}}{6} + \zeta - \sqrt{3}\zeta + 3\zeta^2\right) \\
M_{14} &= \frac{D}{h}(-1+\sqrt{3}) + v\left(-\frac{\sqrt{3}}{6} - \zeta + \sqrt{3}\zeta + 3\zeta^2\right) \\
M_{24} &= -\frac{D}{h}(1+\sqrt{3}) + v\left(\frac{\sqrt{3}}{6} - \zeta - \sqrt{3}\zeta + 3\zeta^2\right).
\end{aligned}
$$

It is easy to see that (13) reduces to (7) when $\zeta = 0$, i.e., when there is no upstream weighting.

As in the previous section, the collocation solution is governed by the parameter $\lambda$. However, now that we have implemented upstream weighting, the value of $\lambda$ now depends on $\zeta$:

$$\lambda = \frac{\beta^2 + 6\beta + 12 + 6\beta\zeta(4 + \beta + \beta\zeta)}{\beta^2 - 6\beta + 12 + 6\beta\zeta(4 - \beta + \beta\zeta)}, \tag{14}$$

which reduces to (10) when $\zeta = 0$.

The solution of (13) is given by

$$q_j = \frac{u_L(\lambda^m - \lambda^j) + u_R(\lambda^j - 1)}{\lambda^m - 1} \tag{15}$$

and

$$r_j = \frac{\rho\lambda^j}{\lambda^m - 1}(u_R - u_L), \tag{16}$$

where

$$\rho = \frac{2\beta m(1 + \beta\zeta)}{\beta^2\zeta^2 + 4\beta\zeta + 2}. \tag{17}$$

Note that (15) and (8) are identical (except for the change in the definition of $\lambda$) and that (16) reduces to (9) when $\zeta = 0$.

Now that $\lambda$ is defined as in (14), it is possible for the collocation solution to oscillate. However, this happens for only a very narrow range of parameter values. Specifically, oscillations occur only when $\beta > 6 + 4\sqrt{3}$ and

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

This region of the $\beta$-$\zeta$ plane may be easily observed. It is the area above the thick dashed curve and below the light dotted line in Figure 2.

We have determined an algorithm for how to choose the value of the upstream parameter $\zeta$ as a function of the Péclet number $\beta$ so as to minimize the difference between (15) and (11) (see [4] for details). This algorithm for the optimal choice of the value of $\zeta$ is given in Table 1 and depicted in Figure 2. (In Table 1, the parameter $\epsilon$ appears. It is a small positive number whose purpose is to ensure that the collocation solution does not oscillate. In all pertinent numerical experiments reported below, its value is set to $10^{-6}$.)

We depict the efficacy of optimal upstream weighting in Figure 3. We see that the collocation solution without upstreaming suffers from significant smearing while the optimal upstream collocation solution is visually indistinguishable from the exact solution.

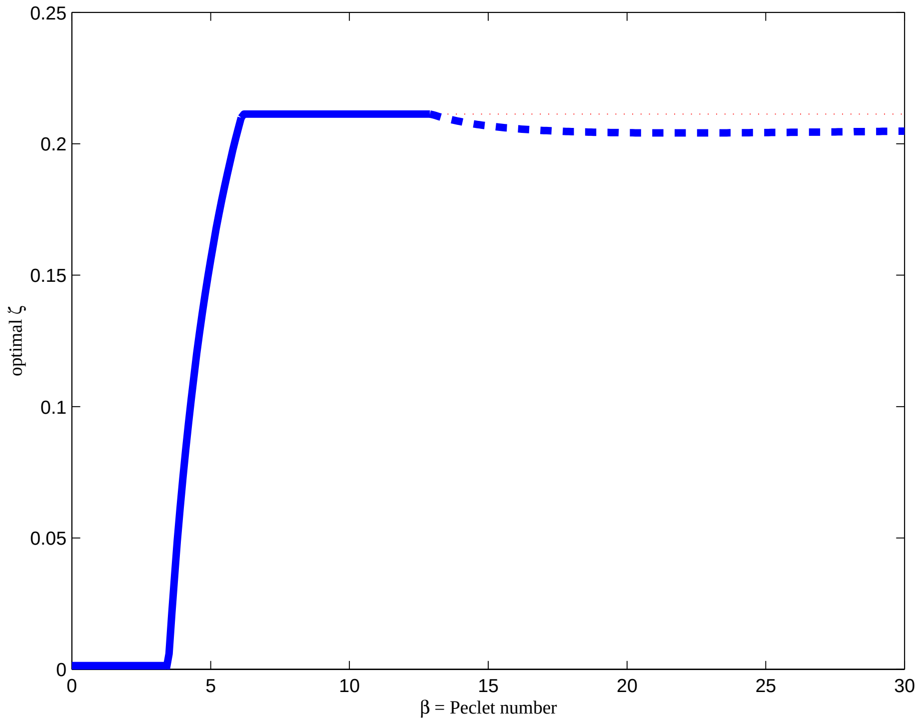


Figure 2: Optimal $\zeta$ as a function of Péclet number $\beta$.

# 4 Optimal upstream weighting with forcing

In this section we build upon the results of the previous one to examine the case where the source/sink term $S(x)$ is no longer zero. As in the previous section, we collocate the second derivative term at the Gauss points and employ upstream weighting for the first derivative term. The question of where to evaluate the forcing function $S(x)$ arises. We shall examine this issue, when $S(x)$ is a linear function, after we give the formulas for the collocation solution of the general problem.

For this case, the matrix equation (13) is replaced by

$$\begin{bmatrix} M_{11} & M_{12} & -M_{11} & M_{14} \\ M_{21} & M_{22} & -M_{21} & M_{24} \end{bmatrix} \begin{bmatrix} q_j \\ r_j \\ q_{j+1} \\ r_{j+1} \end{bmatrix} = \begin{bmatrix} S_{2j} \\ S_{2j+1} \end{bmatrix}, \tag{18}$$

$j = 0, 1, 2, \ldots, m-1$. Here $S_k = S(x_k)$, $k = 0, 1, \ldots, 2m-1$, where these $x_k$'s are the locations at which the forcing function $S(x)$ is evaluated.

As reported in [5], the solution of the matrix equation defined by (18) is

$$q_j = \frac{u_L(\lambda^m - \lambda^j) + u_R(\lambda^j - 1)}{\lambda^m - 1} + \frac{1}{2m} \sum_{k=0}^{2m-1} G_{k,j} S_k \tag{19}$$

Table 1: Optimal $\zeta$ as a function of $\beta$.

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

$$r_j = \frac{\rho\lambda^j}{\lambda^m - 1}(u_R - u_L) + \frac{1}{2m}\sum_{k=0}^{2m-1} G'_{k,j} S_k, \tag{20}$$

where $\lambda$ and $\rho$ are as in (14) and (17), and $S_k$ is as in (18).

The discrete Green's functions $G_{k,j}$ and $G'_{k,j}$ are

$$G_{k,j} = \begin{cases} A_k(\lambda^m - \lambda^j), & k = 0, 1, \ldots, 2j - 1 \\ (C_k - A_k\lambda^m)(\lambda^j - 1), & k = 2j, 2j + 1, \ldots, 2m - 1 \end{cases} \tag{21}$$

and

$$G'_{k,j} = \begin{cases} -\rho A_k \lambda^j, & k = 0, 1, \ldots, 2j - 1 \\ \rho(C_k - A_k\lambda^m)\lambda^j, & k = 2j, 2j + 1, \ldots, 2m - 1, \end{cases} \tag{22}$$

where

$$A_{2\ell} = \frac{(1 + \beta\zeta - \sqrt{3}\beta\zeta^2)\lambda_{\text{num}}\lambda^\ell + \rho_{\text{den}}(-6 - \sqrt{3}\beta - 6\beta\zeta + 6\sqrt{3}\beta\zeta^2)}{v(1 + \beta\zeta)\lambda_{\text{num}}\lambda^\ell(\lambda^m - 1),} \tag{23}$$

$$A_{2\ell+1} = \frac{(1 + \beta\zeta + \sqrt{3}\beta\zeta^2)\lambda_{\text{num}}\lambda^\ell + \rho_{\text{den}}(-6 + \sqrt{3}\beta - 6\beta\zeta - 6\sqrt{3}\beta\zeta^2)}{v(1 + \beta\zeta)\lambda_{\text{num}}\lambda^\ell(\lambda^m - 1),} \tag{24}$$

and

$$C_k = \begin{cases} \dfrac{1 + \beta\zeta - \sqrt{3}\beta\zeta^2}{v(1 + \beta\zeta)}, & k \text{ even} \\[2ex] \dfrac{1 + \beta\zeta + \sqrt{3}\beta\zeta^2}{v(1 + \beta\zeta)}, & k \text{ odd.} \end{cases} \tag{25}$$

The symbols $\lambda_{\text{num}}$ and $\rho_{\text{den}}$ refer respectively to the numerator of (14) and denominator of (17).

Let us observe the relationships between the formulas given above for the collocation solution and those for the exact solution of (1) with the same boundary conditions. We obtain

$$u(x) = u_L + \frac{u_R - u_L}{e^{\beta m} - 1}(e^{\beta m x} - 1) + \int_0^1 G(\xi, x)\, S(\xi)\, d\xi \tag{26}$$

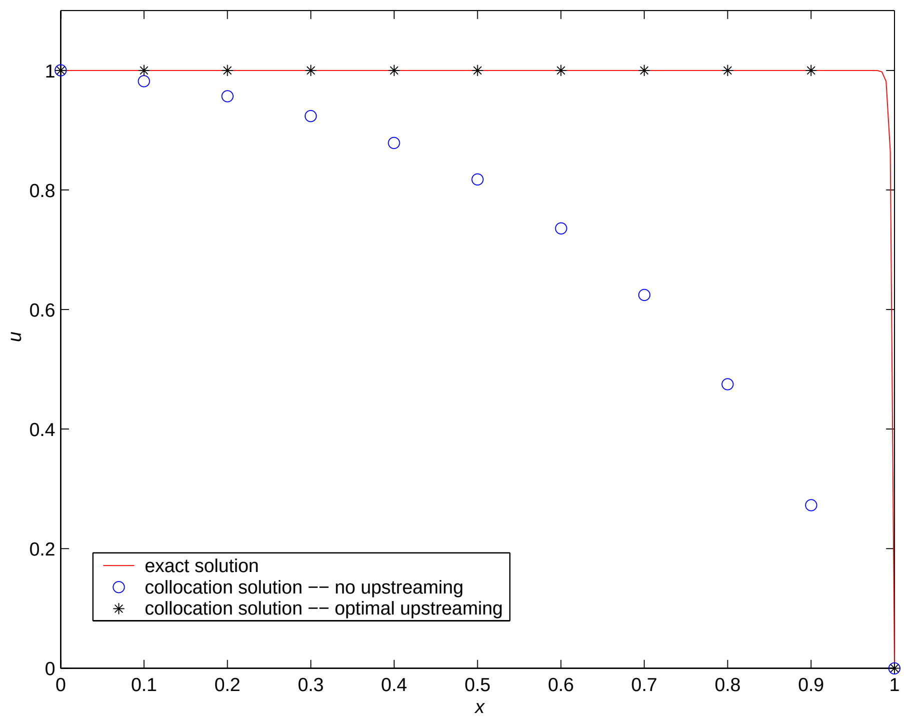


Figure 3: Collocation and exact solutions for $m = 10$, $\beta = 40$, $u_L = 1$, $u_R = 0$, and no forcing.

where the (continuous) Green's function is

$$G(\xi, x) = \begin{cases} \dfrac{\left[e^{\beta m \xi} - 1\right]\left[e^{\beta m(1-\xi)} - e^{\beta m(x-\xi)}\right]}{v(e^{\beta m} - 1)}, & 0 \le \xi \le x \\[2ex] \dfrac{\left[e^{\beta m x} - 1\right]\left[e^{\beta m(1-\xi)} - 1\right]}{v(e^{\beta m} - 1)}, & x \le \xi \le 1. \end{cases}$$

If we evaluate (26) and its derivative at the $j$th mesh point $x_j = \frac{j}{m}$, we obtain

$$u(x_j) = \frac{u_L(e^{\beta m} - e^{\beta j}) + u_R(e^{\beta j} - 1)}{e^{\beta m} - 1} + \int_0^1 G(\xi, x_j)\, S(\xi)\, d\xi \tag{27}$$

where

$$G(\xi, x_j) = \begin{cases} A(\xi)\left[e^{\beta m} - e^{\beta j}\right], & 0 \le \xi \le x_j \\[2ex] \left[\dfrac{1}{v} - A(\xi) e^{\beta m}\right](e^{\beta j} - 1), & x_j \le \xi \le 1 \end{cases} \tag{28}$$

and

$$\frac{du}{dx}(x_j) = \frac{\beta m e^{\beta j}}{e^{\beta m} - 1}(u_R - u_L) + \int_0^1 \frac{\partial G}{\partial x}(\xi, x_j)\; S(\xi)\, d\xi, \tag{29}$$

where

$$\frac{\partial G}{\partial x}(\xi, x_j) = \begin{cases} -A(\xi)\,\beta m\,e^{\beta j}, & 0 \le \xi < x_j \\ \left[\frac{1}{v} - A(\xi)\,e^{\beta m}\right]\,\beta m\,e^{\beta j}, & x_j < \xi \le 1. \end{cases} \tag{30}$$

The function $A(\xi)$ is

$$A(\xi) = \frac{1 - e^{-\beta m \xi}}{v(e^{\beta m} - 1).} \tag{31}$$

Comparison of (19) and (27) reveals that the role of the definite integral in (27) is assumed by the numerical integration rule defined by the summation in (19). It is also evident that the discrete Green's functions (21) and (22) correspond to their counterparts (28) and (30). Continuing with these last comparisons, $C_k$ (see (25)) is the analogue of $\frac{1}{v}$ (and, indeed $C_k = \frac{1}{v}$ when $\zeta = 0$). Finally, (23) and (24) clearly correspond to (31), but this relationship is more tenuous than those discussed above due to the fact that $\xi$ is a "dummy" variable and thus has no direct connection to the differential equation (1) or its upstream collocation discretization.

For the specific case of linear forcing, we set $S(x) = \alpha_0 + \alpha_1 x$ and we evaluate this function at the points

$$\left(\frac{2p+1}{2} + \psi \pm \frac{1}{\sqrt{12}}\right) h, \tag{32}$$

$p = 0, 1, \ldots, m-1$. If $\psi = 0$, then we are evaluating $S(x)$ at the Gauss points. It is clear that we must restrict $\psi$ to lie in the interval $\left[\frac{1}{\sqrt{12}} - \frac{1}{2}, \frac{1}{2} - \frac{1}{\sqrt{12}}\right]$. In this paradigm, many of the formulas above are greatly simplified (see [5] for details). In particular, we obtain

$$u(x_j) = \gamma_0\alpha_0 + \gamma_1\alpha_1 + \gamma_2, \tag{33}$$

where

$$\gamma_0 = \frac{1}{v}\left(\frac{j}{m} - \frac{e^{\beta j} - 1}{e^{\beta m} - 1}\right), \tag{34}$$

$$\gamma_1 = \frac{1}{2\beta m v}\left[(\beta j + 2)\frac{j}{m} - (\beta m + 2)\frac{e^{\beta j} - 1}{e^{\beta m} - 1}\right], \tag{35}$$

and

$$\gamma_2 = \frac{u_L(e^{\beta m} - e^{\beta j}) + u_R(e^{\beta j} - 1)}{e^{\beta m} - 1} \tag{36}$$

The upstream collocation solution of the same problem is

$$q_j = \delta_0\alpha_0 + \delta_1\alpha_1 + \delta_2, \tag{37}$$

where

$$\delta_0 = \frac{1}{v}\left(\frac{j}{m} - \frac{\lambda^j - 1}{\lambda^m - 1}\right), \tag{38}$$

$$\delta_1 = \frac{1}{2\beta m v}\left\{[\beta j + 2(\beta\psi + \beta\zeta + 1)]\,\frac{j}{m} - [\beta m + 2(\beta\psi + \beta\zeta + 1)]\,\frac{\lambda^j - 1}{\lambda^m - 1}\right\}, \quad (39)$$

and

$$\delta_2 = \frac{u_L(\lambda^m - \lambda^j) + u_R(\lambda^j - 1)}{\lambda^m - 1} \quad (40)$$

To minimize the difference between (33) and (37), we need to respectively compare (34), (35), and (36) to (38), (39), and (40). In the previous section, we optimally minimized the maximum value of

$$\left|\frac{e^{\beta j} - 1}{e^{\beta m} - 1} - \frac{\lambda^j - 1}{\lambda^m - 1}\right|,$$

$j = 1, 2, \ldots, m-1$, so we need compare now only (35) and (39). The difference between these is minimized by setting

$$\psi = -\zeta. \quad (41)$$

That is, we will evaluate the forcing function $S(x)$ at precisely the same optimal upstream locations at which we evaluate the term $\widehat{\frac{du}{dx}}$.

We illustrate the efficacy of this choice in Figure 4. The Péclet number is $\beta = 50$, and the forcing function is $S(x) = -10 + 200x$. We solved the problem many times, using $\psi = -0.20, -0.18, -0.16, \ldots, 0.20$ and the value given in (41). We see that using optimal upstream weighting in conjunction with (41) again yields extremely accurate results. Furthermore, as $\psi$ moves farther from its optimal value given in (41), the approximations become less accurate.

# 5 Summary and conclusions

Our recent work concerning analytical formulas for the solution of the Hermite collocation discretization of the one-dimensional convection-diffusion equation has been summarized in this paper. Due to space constraints, the exposition here is lacking in detail, which the interested reader may find in [3], [4], and [5]. The unifying concept of these works is that if one applies "optimal upstream weighting" to convection-dominated problems, then one obtains numerical solutions that don't suffer from the "smearing" effect, are non-oscillatory, and are visually indistinguishable from the corresponding exact solution.

# References


[1] Allen, M. B. "How Upstream Collocation Works." *Int. J. Num. Meth. Eng.*, 19 (1983): 1753–1763.

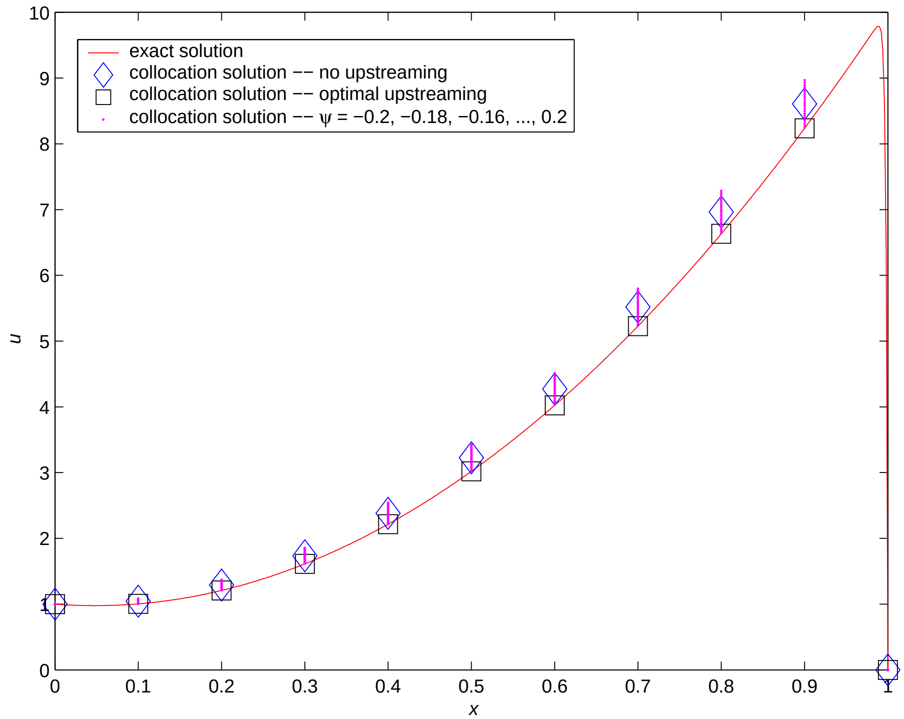


Figure 4: Exact and collocation solutions for $m = 10$, $\beta = 50$, $S(x) = -10+200x$, $u_L = 1$, $u_R = 0$.


[2] Brill, S. H. "Analytical Solution of Hermite Collocation Discretization of Self-Adjoint Ordinary Differential Equations." *International Journal of Differential Equations and Applications*, to appear.

[3] Brill, S. H. "Analytical Solution of Hermite Collocation Discretization of the Steady-State Convection-Diffusion Equation." *International Journal of Differential Equations and Applications* 4.2 (2002): 141–155.

[4] Brill, S. H. "Optimal Upstream Collocation Solution of the One-Dimensional Steady-State Convection-Diffusion Equation." *International Journal of Applied Mathematics*, to appear.

[5] Brill, S. H. "Analytical Upstream Collocation Solution of a Forced One-Dimensional Steady-State Convection-Diffusion Equation." Submitted to *BIT Numerical Mathematics.*

[6] P. M. Prenter, *Splines and Variational Methods*, John Wiley & Sons, New York (1975).