Reduced order models for diffusion systems using singular perturbations

Bharath Bhikkaji and Torsten Söderström

Department of Systems and Control, Information Technology, Uppsala University, P O Box 27, SE 75121 Uppsala, Sweden, email bb@SysCon.uu.se, ts@SysCon.uu.se

October 16, 2000

Abstract

In this paper, we consider a special case of the one dimensional heat diffusion across a homogeneous wall. This physical system is modeled by a linear partial differential equation, which can be thought of as an infinite dimensional dynamic system. To simulate this physical system, one has to approximate the underlying infinite order system by a finite order approximation. In this paper we first construct a simple and straightforward approximate finite order model for the true system. The proposed approximate models may require large model order to approximate the true system dynamics in the high frequency regions. To avoid the usage of higher order models, we use a scheme similar to singular perturbations to further reduce the model order.

Keywords: Diffusion, State-space model, Singular perturbations.
1 Introduction

A lot of physical process, like flow of heat across a temperature gradient, flow of chemicals across a concentration gradient are examples of diffusion phenomena. Systems which involve diffusion phenomena are called diffusion systems, see [3], [9]. Diffusion systems are normally modeled by a linear partial differential equations (LPDE). LPDE’s can be thought of as infinite dimensional linear dynamic systems. To simulate a diffusion system, one has to approximate the underlying infinite dimensional dynamical system by a finite order model. In some cases, in order to achieve certain realistic error bounds in the simulation of the system, one would be forced to use an approximate model of large model order.

Singular perturbation is a technique to reduce the model order of a large scale dynamic systems (dynamic systems with large model order), which involves interacting dynamic phenomena of varying time constants. The theme behind singular perturbation is that, one assumes the given large scale system can be characterized by two modes, a fast mode and a slow mode. Model reduction is achieved by neglecting the fast phenomena or the fast mode, and later, the reduced model is improved by adding certain correction factors. For more details refer to [8], [5] and [7].

In this paper, we consider the simple case of a heat diffusion in a homogeneous wall. We first show that the LPDE, that results from this problem, is actually an infinite order dynamical system and hence needs model reduction. We propose certain finite order approximations for the given infinite order dynamical system, and these finite order models occur as natural choice for the true system. The proposed approximations may require large model orders to match the true dynamics in the high frequency regions. To overcome this problem of using a higher order model, we use a scheme, which is similar to singular perturbations approach, to further reduce the model order without compromising much on the error in the high frequency regions.

This paper is organized into five sections. In Section 2 we define and extensively discuss the problem to be solved and also discuss our approach to solve the problem. In Section 3 we derive the true infinite order dynamics of the system, while in Section 4 we construct an approximate finite order model for the infinite order system and discuss about them. In Section 5 we reduce higher order models to obtain lower order models that have less error in high frequency regions. In Section 6 we numerically analyze the convergence of the proposed approximate models.
2 A case study: the heat diffusion equation

2.1 Statement of the problem

Consider a homogeneous wall of thickness d units. Let x be the coordinate across the wall, $T_e$ the temperature on the exterior side, $T_i$ the temperature on the interior side and $q_i$ the supplied heat on the interior side, cf Figure 1.

Let $T = T(x, t)$ denote the temperature at position $x$ and at time $t$. Then the heat flow and the temperature are described by

$$\frac{\partial T(x, t)}{\partial t} = \alpha \frac{\partial^2 T(x, t)}{\partial x^2}$$

with boundary constraints

$$q_k(t) = -\kappa \left. \frac{\partial T(x, t)}{\partial x} \right|_{x=0},$$

$$T_e(t) = T(d, t)$$

For simplicity we consider the intitial condition $T(x, 0) = 0$. In the model (1)-(3) the diffusion coefficient $\alpha$ and and the thermal conductivity $\kappa$ are parameters that depend on the material of the wall.

Needless to say, both the case selected for study and its geometry are simple. While the geometry is chosen for technical ease, the simple case is aimed at getting insight into more difficult systems for future work.

The problem to be solved is, given the parameters $\alpha$ and $\kappa$ and the boundary data \{q_k(kh), T_e(kh), \}_k^{N}, for a particular given sampling interval of h units, estimate the interior temperature $\{T_i(kh)\}_k^{N}$.

Note that, it has been inherently assumed that the thermal properties of the wall are not affected by the variations of external temperature or the heat supplied. This can also be seen from the model, (1)-(3), as the parameters $\alpha$ and $\kappa$ are assumed to be constants.

Since the wall is independent of the boundary data \{q_k(kh), T_e(kh), \}_k^{N}, one can think of the wall as a continuous time dynamic system, see [12], with $u(t) =$
\((q_i(t), T_e(t))^T\) as the input and \(y(t) = T_i(t)\) as the output, see Figure 2. The impulse response \(h(t)\) of the wall or its Laplace transform, the transfer function \(H(s)\) is determined from the governing equations \((1)-(3)\). Once the transfer function \(H(s)\) or the impulse response \(h(t)\) is determined, \(T_i(t)\) can be computed either as a convolution of the input \(u(t)\) with the impulse response \(h(t)\) or as a inverse Laplace transform of the product of \(U(s)\) and \(H(s)\), where \(U(s)\) is the Laplace transform of the input \(u(t)\).

In the next section we determine the transfer function \(H(s)\) of the wall and show that it is a transcendental function, or in system’s terminology, it is of infinite order. Since the transfer function \(H(s)\) of the wall is of infinite order, one has to approximate \(H(s)\) by a finite order model, which is referred to as model reduction in system’s terminology. Hence, we propose certain finite order transfer functions, which arise as a natural choice of approximation for the infinite order transfer function \(H(s)\). It turns out that one needs to use a large model order to approximate \(H(s)\) more closely in the high frequency regions. To overcome this problem of using a higher order approximation, we use a model reduction technique, which is similar singular perturbations to reduce these higher order models into lower order ones. These models approximate \(H(s)\) better in the high frequency regions than those corresponding lower order models directly used.

We strike a note of caution here when we say that, ”the lower order models obtained by reducing the higher order models, approximate \(H(s)\) better in the high frequency regions than those corresponding lower order models directly used”. It would be more appropriate to say that the lower order models, obtained by using singular perturbations, approximate \(H(s)\) better in certain desired high frequency regions and not in general throughout the high frequency regions.

It is well known that there are a lot of well established numerical schemes, like Galerkin, Finite difference and Collocation, see [4], that can be used to compute \(T_i(t)\) from \((1)-(3)\). These numerical schemes, when applied to \((1)-(3)\), can also be interpreted as schemes which construct approximate transfer functions for the true transfer function \(H(s)\), see [11] and [1]. Even though Galerkin, Finite
difference and Collocation have a well-established theory for their convergence, they are very general in their approach and do not exploit the structure of the given problem. In this paper we take a systems approach rather than a numerical analysis approach and propose certain models, which exploits the structure of the given problem. These models turn out to approximate the true system better than the Galerkin, Finite difference and Collocation, for a given model order.

3 The system dynamics

The defined system is linear and has two inputs, $q_t$ and $T_e$, and one output, $T_i$. Interestingly enough, one can rather easily derive the transfer function of the system, which will give an input-output description of the dynamics, [11]. For convenience we repeat the derivation here.

First take the Laplace transform of (1) which gives

$$sT(x, s) = \alpha \frac{d^2}{dx^2} T(x, s)$$

(4)

where $T(x, s)$ is the Laplace transform of the temperature $T(x, t)$. Regarding $s$ as a fixed parameter, (4) is a standard ordinary differential equation. It is found that the general solution to (4) can be written as

$$T(x, s) = A(s)\cosh(x\sqrt{\frac{s}{\alpha}}) + B(s)\sinh(x\sqrt{\frac{s}{\alpha}})$$

(5)

Further, taking Laplace transforms on (2),(3) and evaluating the constants $A(s)$ and $B(s)$ gives the following expression for temperature at the interior side.

$$T_i(s) = \frac{\tanh(d\sqrt{\frac{s}{\alpha}})}{\kappa \sqrt{\frac{s}{\alpha}}} q_t(s) + \frac{1}{\cosh(d\sqrt{\frac{s}{\alpha}})} T_e(s).$$

(6)

$$G_1(s) \quad G_2(s)$$

Note that both $G_1(s)$ and $G_2(s)$ are bounded transcendental functions, with the same set of poles

$$P_n = -\frac{\pi^2}{d^2}(n - \frac{1}{2})^2\alpha, \quad n = 1, 2, \ldots,$$

(7)

and only $G_1(s)$ has zeros,

$$Z_n = -\frac{\pi^2}{d^2}n^2\alpha \quad n = 1, 2, \ldots.$$  

(8)

The static gains of $G_1(s)$ and $G_2(s)$ are

$$G_1(0) = \frac{d}{\kappa}, \quad G_2(0) = 1.$$  

(9)
Since all the poles of \( G_1(s) \) and \( G_2(s) \), (7), are real and negative, the system is stable and the transfer functions \( G_1(s) \) and \( G_2(s) \) are bounded. Boundedness of \( G_1(s) \) and \( G_2(s) \) implies that, the output \( T_i(t) \) depends continuously on the input data \( q_i(t) \) and \( T_e(t) \). Continuous dependence along with the uniqueness of the solution renders the problem, (1)-(3), well posed in numerical analysis terminology.

It is worth noting that both \( G_1(s) \) and \( G_2(s) \) are of low pass character, see Figures 3, 4 and 5. Hence the approximate models should also be of low pass character with the same static gain. Otherwise the approximate models would not reach the steady state as \( t \to \infty \) in the time domain. The low pass character of \( G_1(s) \) and \( G_2(s) \) would also enable us to choose the approximate model depending on the frequency contents of the input signals.

If one takes a closer look at the Figures 3, 4 and 5, one can see that \( |G_2(i\omega)| \to 0 \) at a rate much faster than \( |G_1(i\omega)| \), as \( \omega \to \infty \). The faster convergence of \( |G_2(i\omega)| \to 0 \) is apparently due to the fact it has no zeros and it is an all pole system. Due to its fast convergence \( |G_2(i\omega)| \) may not need a higher order approximation, as \( |G_2(i\omega)| \) is almost close to 0 in the high frequency regions. In the case of \( |G_1(i\omega)| \), it behaves more like \( \frac{\kappa}{\alpha^2} \) in the high frequency regions, as \( \tanh(d\sqrt{\frac{\omega}{\alpha}}) \approx 1 \) for a large \( s = i\omega \). Hence one may have to use a higher order models to approximate \( G_2(s) \) well in the high frequency regions.

Finally, we look at the effect of the parameters \( \alpha \) and \( \kappa \) on the transfer functions \( G_1(s) \) and \( G_2(s) \). From (6) one can see that the parameter \( \alpha \) merely re-scales the frequency axis, as \( G_1(s) \) and \( G_2(s) \) can also be considered as functions of \( \frac{1}{\alpha} \). The parameter \( \kappa \) just scales the magnitude of \( G_1(s) \) by \( \kappa \) and has no effect on \( G_2(s) \). To make our analysis more general, we scale the frequency axis by using the transformation \( s \mapsto \alpha s \), to discount the effect of the parameter \( \alpha \) in (6). For a given \( \alpha \), one can always revert back to the original frequency domain by using the inverse transformation \( s \mapsto \frac{1}{\alpha} s \) on the scaled frequency domain.

### 4 Approximate models

Without much ado, we directly introduce the approximate models. Since \( G_1(s) \) is characterized by the poles (7) and the zeros (8), a natural candidate for approximating \( G_1(s) \) is

\[
G_1^n(s) = K_1 \frac{(s - z_1)(s - z_2) \ldots (s - z_{n-1})}{(s - p_1)(s - p_2) \ldots (s - p_n)},
\]

where

\[
K_1 = \frac{\frac{p_1 p_2 \ldots p_n}{z_1 z_2 \ldots z_{n-1}}}{d}
\]

6
is the normalizing constant which renders

\[ G_1(0) = G_1^n(0). \] (12)

Further, \( p_i \) and \( z_i, i = 1, 2, \ldots, n \) are the poles and zeros, respectively, of the true transfer functions with their frequencies scaled by \( \alpha \), that is

\[ p_i = \frac{P_i}{\alpha} \] (13)

and

\[ z_i = \frac{Z_i}{\alpha} \] (14)

where \( P_i, i = 1, 2, \ldots, n \) are as defined in (7) and \( Z_i, i = 1, 2, \ldots, n - 1 \) are as defined in (8).

Following a similar pattern as above and noting that \( G_2(s) \) has no zeros, we define

\[ G_2^n(s) = K_2 \frac{1}{(s - p_1)(s - p_2)\ldots(s - p_n)}, \] (15)

where

\[ K_2 = (-1)^n p_1 p_2 \ldots p_n \] (16)

is the normalizing constant which renders

\[ G_2(0) = G_2^n(0) \] (17)

and \( p_i, i = 1, 2, \ldots, n \) are as defined in (13).

One can see from the Figures 3, 4 and 5, that the approximate transfer functions \( G_1^n(s) \) and \( G_2^n(s) \) track the true true transfer functions, \( G_1(s) \) and \( G_2(s) \), upto a certain frequency but deviate after that. In fact one can say that \( G_1(s) \) and \( G_2(s) \) are well approximated by \( G_1^n(s) \) and \( G_2^n(s) \) at the low frequency regions, but the approximation deteriorates as the frequency increases. Note that by construction both \( |G_1^n(i\omega)| \) and \( |G_2^n(i\omega)| \rightarrow 0 \) as \( \omega \rightarrow \infty \). Hence one would expect the error \( |G_i(i\omega) - G_i^n(i\omega)| \) to increase from zero at \( \omega = 0 \) and attain a maximum at some finite \( \omega \) and then decrease to zero as \( \omega \rightarrow \infty \). This can also be seen in the Figures 6 and 7.

To check the consistency of the proposed approximate models \( G_1^n(s) \) and \( G_2^n(s) \), we plot, in Figure 8, \( \max_\omega |G_i(i\omega) - G_i^n(i\omega)| \) for different model orders. Note from Figure 8 that \( \max_\omega |G_i(i\omega) - G_i^n(i\omega)| \) decreases as \( n \) increases. From Figure 8 we further conclude that \( \max_\omega |G_i(i\omega) - G_i^n(i\omega)| \rightarrow 0 \) as \( n \rightarrow \infty \). Hence we also conclude that \( G_i^n(s) \rightarrow G_i(s), i = 1, 2. \)
Table 1: $\alpha$ in $m^2/hr$ for typical building materials.

<table>
<thead>
<tr>
<th>Material</th>
<th>$\alpha$ in $m^2/hr$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aluminium</td>
<td>0.3383</td>
</tr>
<tr>
<td>Concrete</td>
<td>0.0037</td>
</tr>
<tr>
<td>Wood</td>
<td>$3.1579 \times 10^{-4}$ - $6.3158 \times 10^{-4}$</td>
</tr>
<tr>
<td>Mineral wool</td>
<td>$4.2857 \times 10^{-5}$ - $2.5714 \times 10^{-4}$</td>
</tr>
</tbody>
</table>

One has to be very careful, when one speaks about about high and low frequency regions, as both $G_1(s)$ and $G_2(s)$, in the Figures 3, 4 and 5, are plotted for normalised frequencies. Table 1 the values of $\alpha$, in $m^2/hr$, for certain typical building materials. Hence a normalised frequency of $10^4$ rads/hr would correspond to 37 rads/hr for concrete and 3383 rads/hr for aluminum. What is a high frequency for a particular material may not be high for some other materials. Nevertheless, when we speak about frequencies with respect to the transfer functions, we mean only normalised frequencies.

Since $|G_1(s)| \rightarrow 0$ at a very slow rate, at $\frac{1}{s^2}$ as $s \rightarrow \infty$, it has considerable magnitude even at the high frequencies and hence needs to be well approximated in the high frequency regions. Note that, form the Figures 3, 4 and 5, $|G_1(s)| \rightarrow 0$ at a faster rate than $|G_1(s)|$. This is due to fact that $G_1^n(s)$ has a pole excess, (10), so $G_1^n(s) \sim \frac{K}{s}$ for a large $s$. This faster convergence deteriorates the approximation as the frequency increases. If possible one would like to have an improved approximation in the high frequency regions with a lower order model.

In the case of $G_2(s)$, things are much simpler, since $|G_2(s)| \rightarrow 0$ at an exponential rate, see Figures 3, 4 and 5, which is due to the fact that $G_2(s)$ has no zeros to nullify the effect of the poles. Note that, from the Figures 3, 4 and 5, both $|G_2(s)|$ and $|G_2^n(s)|$ are very close to zero for frequencies above 1000 rads/hr. Hence improvement of the approximate transfer function $G_2^n(s)$, with a low model order is necessary only if the input $T_e(t)$ has high frequency components of very high magnitude.

As we have mentioned before, the choice of the model order depends very much on the spectral contents of the input signals. If one restricts oneself to a scenario of using the proposed model for the materials used for building houses, then the input $T_e(t)$ or the external temperature, is normally constructed from the Test Reference Year Data (TRY data), see [2], available at that place where the approximate model is going to be used. TRY Data, at any given place, consists of hourly samples of the temperature over a period of one year with a minimum difference of 0.1°C between two non-identical samples of the temperature. TRY data normally do not have high frequency components of considerable magnitude. This is apparently due to the fact that, temperature at any given place does not
show marked differences in short intervals of time. One would have to view
them in large windows of time like months or seasons to see large variations in
temperature. Hence even for a frequency of 10 rads/hr, is too high a frequency
value for the external temperature \( T_e(t) \), to have considerable magnitude. Thus
a lower order model, say \( n = 7 \) or 8 or 9 may suffice for \( G_2(s) \), due to the low
frequency contents of the input, in the case of materials used for building houses.
However, if one wants to use the proposed model, in situations where the input
\( T_e(t) \) has high frequency components, then one has no choice, but to use a higher
order model depending on the frequency contents of the input.

The input \( q_i(t) \) may or may not contain high frequency components, depending
on the application, for which the approximate model, \( G_1^n(s) \), is going to be used. In most practical applications involving building materials, \( q_i(t) \) is used as
a control signal to keep the internal temperature, \( T_i(t) \), constant or close to a
desired value. In such cases \( q_i(t) \) would mostly bootstrap \( T_e(t) \) and would have
frequency contents almost similar to that of \( T_e(t) \). In such cases again one may
not need very high model order, possibly a model of order \( n = 7 \) or 8 or 9 would
serve the purpose. But if the application involves usage of \( q_i(t) \), which has high
frequency components with considerable magnitude then one has no other option
but to use a higher order model.

As we seen before, from Figures 3, 4 and 5, the approximations \( G_1^n(s) \) and
\( G_2^n(s) \) deteriorate as the frequency increases. The primary reason for the poor
performance of \( G_1^n(s) \) and \( G_2^n(s) \) at the high frequencies is due the non inclusion
of the higher order poles and zeros in the lower order models. For any general
transfer function, the lower order poles and zeros (poles and zeros close to the
origin) play a larger role in determining the magnitude of transfer function in the
low frequency regions, whereas the higher order poles and zeros (poles and zeros
that are quite far away from the origin) have negligible effect on the magnitude
of transfer functions at the low frequency regions. But as the frequency increases
the higher order poles and zeros start playing a greater role in determining the
magnitude of the transfer functions, as \( \omega \rightarrow \infty \) almost all poles have equal effect.
Hence to get better approximations for \( G_1(s) \) and \( G_2(s) \) in the high frequency
regions, one would have to incorporate the higher order poles and zeros in (10)
and (15). In effect one has to increase the model order for better approximations
in the high frequency regions.

Since the inputs \( q_i(t) \) and \( T_e(t) \) are band limited signals there may arise a lot
of situations, where the spectral densities of \( q_i(t) \) and \( T_e(t) \), may decay rapidly
to zero after a particular frequency. A typical example would be, the spectral
densities of \( q_i(t) \) and \( T_e(t) \) having the shape of a low pass filter or a band pass
filter. In such situations, the following model reduction technique may come in
handy.
Figure 3: The approximate transfer functions (10) and (15) along with their respective true transfer functions $G_1(s)$ and $G_2(s)$ with the model order $n = 3$. The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line. The frequency axis is scaled by the parameter $\alpha$ for both $G_1(s)$ and $G_2(s)$ and the magnitude of $G_1(s)$ is scaled by $\kappa$. 
Figure 4: The approximate transfer functions (10) and (15) along with their respective true transfer functions $G_1(s)$ and $G_2(s)$, with the model order ($n = 5$). The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line. The frequency axis is scaled by the parameter $\alpha$ for both $G_1(s)$ and $G_2(s)$ and the magnitude of $G_1(s)$ is scaled by $\kappa$. 
Figure 5: The approximate transfer functions (10) and (15) along with their respective true transfer functions $G_1(s)$ and $G_2(s)$, with the model order ($n = 7$). The true transfer function (of infinite order) is plotted with a dotted line and the approximate transfer function (of finite order) is plotted with a solid line. The frequency axis is scaled by the parameter $\alpha$ for both $G_1(s)$ and $G_2(s)$ and the magnitude of $G_1(s)$ is scaled by $\kappa$. 
Figure 6: The error, $|G_1(\omega) - G^n_1(\omega)|$, in the approximate transfer function, $G^n_1(\omega)$, for model orders $n = 3, 5, 7$. The frequency axis is scaled by the parameter $\alpha$ and the magnitude of $G_1(s)$ and $G^n_1(s)$ is scaled by $\kappa$. 
Figure 7: The error, $|G_2(i\omega) - G_2^n(i\omega)|$, in the approximate transfer function, $G_2^n(i\omega)$, for model orders $n = 3, 5, 7$. The frequency axis is scaled by the parameter $\alpha$. 
Figure 8: The maximum error, $\max_{\omega} | G_1(i\omega) - G_1^n(i\omega) |$ and $\max_{\omega} | G_2(i\omega) - G_2^n(i\omega) |$, in the approximate transfer functions $G_1^n(i\omega)$ and $G_2^n(i\omega)$ respectively. The frequency axis is scaled by the parameter $\alpha$ and the magnitude of $G_1(s)$ and $G_1^n(s)$ is scaled by $\kappa$. 
5 Model reduction using singular perturbation

As we have mentioned in the previous section, the poor performance of the lower order approximations is due to the non-inclusion of the higher order poles and zeros. By using a lower order model and thereby not including the higher order poles and zeros, one cannot achieve a good approximation in the high frequency regions. To have better approximations there one must include the higher order poles and zeros and thereby use a higher order model. As a trade off, we device a model reduction scheme, which is similar to the singular perturbation technique used for reducing the model order of stiff linear systems, see [8],[5] and [7]. In this scheme we take a model of large order and then reduce it to lower order model without completely neglecting the higher order poles and zeros. The lower order model thus obtained by reducing the higher order model, tracks the true transfer function for a longer frequency window than the corresponding lower order transfer functions obtained directly by using (10) and (15). This is not without a cost, as these approximate models do not converge to zero as the frequency tends to infinity. First, we rewrite (10) and (15) using partial fraction decompositions as

\[
G_1^n(s) = K_1 \left[ \frac{A_1}{s - p_1} + \frac{A_2}{s - p_2} + \ldots + \frac{A_n}{s - p_n} \right] \quad (18)
\]

\[
G_2^n(s) = K_2 \left[ \frac{B_1}{s - p_1} + \frac{B_2}{s - p_2} + \ldots + \frac{B_n}{s - p_n} \right] \quad (19)
\]

where

\[
A_i = \frac{(z_1 - p_i)(z_2 - p_i)(z_3 - p_i) \ldots (z_{n-1} - p_i)}{(p_1 - p_i)(p_2 - p_i) \ldots (p_{i-1} - p_i)(p_{i+1} - p_i) \ldots (p_n - p_i)}. \quad (20)
\]

and

\[
B_i = \frac{1}{(p_1 - p_i)(p_2 - p_i) \ldots (p_{i-1} - p_i)(p_{i+1} - p_i) \ldots (p_n - p_i)}. \quad (21)
\]

A justification for the expressions (20) and (21) is provided in the appendix.

Let us first consider \( G_1^n(s) \). As a first step towards model reduction we separate the given transfer function \( G_1^n(s) \) into fast and slow modes.

\[
G_1^n(s) = \hat{G}_1^m(s) + H_1(s) \quad (22)
\]

where

\[
\hat{G}_1^m(s) = K_1 \left[ \frac{A_1}{s - p_1} + \frac{A_2}{s - p_2} + \ldots + \frac{A_m}{s - p_m} \right], \quad (23)
\]

\[
H_1(s) = K_1 \left[ \frac{A_{m+1}}{s - p_{m+1}} + \ldots + \frac{A_n}{s - p_n} \right] \quad (24)
\]
Note that $\hat{G}_1^n(s)$ corresponds to the slow component or slow phenomena, as all its poles are dominating the poles of $H_1(s)$. Vice versa, $H_1(s)$ corresponds to the fast component or fast phenomena, as all its poles are relatively fast when compared to the poles of $H_1(s)$. Looking at it in the time domain by taking inverse Laplace transform of (23) and (24), one can see that the inverse Laplace transform of both (23) and (24) are linear combinations of exponentials with the inverse Laplace transform of (24) converging to zero at a rate faster than the inverse Laplace transform of (23).

As in singular perturbations, we first obtain an intermediate lower order model by neglecting the fast component, that is, we obtain an intermediate reduced order model $\hat{G}_1^n(s)$, by dropping $H_1(s)$ in (22), and then add some correction factors. These correction factors are obtained by approximating the fast component, $H_1(s)$, by its static gain $H_1(0)$, and thus we have the reduced order model

$$G_{R1}^m = \hat{G}_1^n(s) + H_1(0)$$  (25)

for (22).

Similarly by rewriting (19), as a sum of fast modes and slow modes and approximating the fast mode by its static gain, we get a reduced order model

$$G_{R2}^m = \hat{G}_2^n(s) + H_2(0),$$  (26)

for $G_2^n(s)$.

If one looks at the approximate models (25) and (26) in the time domain, one can see that the fast components $H_1(s)$ and $H_2(s)$, which correspond to

$$K_1 \sum_{i=m+1}^{n} A_i \exp(-p_i t)$$  (27)

and

$$K_2 \sum_{i=m+1}^{n} B_i \exp(-p_i t),$$  (28)

respectively, in the time domain, have been approximated by

$$K_1 \sum_{i=m+1}^{n} \frac{A_i}{-p_i} \delta(t)$$  (29)

and

$$K_2 \sum_{i=m+1}^{n} \frac{B_i}{-p_i} \delta(t),$$  (30)
respectively. To put it in words, the quickly decaying exponentials have been approximated by delta functions or impulse functions. This kind of an approximation has been extensively used in fluid dynamics for a long time, see pages 481 - 519 in [6].

In Figures 9, 10 and 11, we have plotted $\hat{G}_R^m(s)$ along with $G_i^m(s)$ and true transfer functions $G_i(s)$, for $i = 1, 2$. One can see that $\hat{G}_R^m(s)$ tracks $G_i(s)$, for a longer period (for a larger frequency window) than $G_i^m(s)$, but $\hat{G}_R^m(s)$ does not converge to zero as $s \to \infty$. This is due to the fact that

$$\lim_{s \to \infty} \hat{G}_R^m(s) = H_i(0). \quad (31)$$

As we have mentioned before, this model reduction scheme comes handy only when the spectral density of, $q_i(t)$ and $T_e(t)$, decays quickly after a particular frequency. Since one can use the extra region of approximation to capture the tail of the spectrum of $q_i(t)$ and $T_e(t)$ and avoid a few extra computations. To make it clear, consider a case where the normalised frequency contents of $q_i(t)$ are concentrated within the frequency window of $[0, 100]$ and decay very rapidly after 100 $\text{rads/hr}$, then instead of using a higher order model of $n = 7$ or 8 or 9, once can use a reduced order model with $m = 5$ and $n = 11$, that is 5 slow components and 6 fast components, see Figures 10, and thereby save a few extra computations.

6 Concluding discussion

We conclude this paper by discussing a few questions, which we think are important. In section 1, we claimed that the models proposed, (10) and (15), in this paper are better than the models obtained by using conventional schemes like Finite difference, Chebyshev Collocation, Chebyshev Galerkin and Chebyshev Tau. In [1], we have shown by simulations, that for a given model order $n$, Chebyshev Collocation gives a better approximation for this problem, when compared to Chebyshev Galerkin and Chebyshev Tau. From [11] one can also conclude that for a given model order $n$ Chebyshev Collocation gives a better approximation, for this problem, than Finite difference approximation. Hence it is enough if we compare the models proposed in this paper with Chebyshev Collocation. In Figures 12, 13, 14, 15, 16 and 17 we have plotted the true transfer functions along with their respective Chebyshev Collocation approximation and the proposed models (10) and (15). In the case of $G_2(s)$, from the Figures 15, 16 and 17, we can see that the proposed model, $G_2^m(s)$, approximates $G_2(s)$ much better than a Chebyshev Collocation model of same order. In the case of $G_1(s)$, one cannot infer from the Figures 12, 13, 14, which one is better. In the low frequency regions both, $G_1^m(s)$ and the Chebyshev Collocation approximation, seems to approximate $G_1(s)$ quite well. Since $G_1^m(s) \to 0$, one can say that $G_1^m(s)$ approximates $G_1(s)$ better, in the high frequency regions, than the Chebyshev
Figure 9: The approximate transfer function (10), plotted in solid line, along with the true transfer function in the dotted line and the singularly perturbed model (25) in dashed dots. The approximate transfer function, (10), in the figure, is plotted with $n = 3$ and singularly perturbed model (25) is plotted with $m = 3$ and $n = 10$. 
Figure 10: The approximate transfer function (25), plotted in solid line, along with the true transfer function in the dotted line and the singularly perturbed model (25) in dashed dots. The approximate transfer function, (10), in the figure, is plotted with $n = 5$ and singularly perturbed model (25) is plotted with $m = 5$ and $n = 10$. 
Figure 11: The approximate transfer function (10), plotted in solid line, along with the true transfer function in the dotted line and the singularly perturbed model (25) in dashed dots. The approximate transfer function, (10), in the figure, is plotted with $n = 7$ and singularly perturbed model (25) is plotted with $m = 7$ and $n = 15$. 

21
Collocation approximation, as the Chebyshev Collocation approximation does not tend to zero as the frequency tends to infinity. But in the grey areas between the high and the low frequency region (the interval \([100, 1000]\) rads/hr in the Figures 12, 13 and 14) it is not clear, which is better.

If one compares the singularly perturbed reduced order models \(G^n_{R1}\), (25), with a Chebyshev Collocation approximation, see Figures 18, 19, 20, it is fairly clear that the singularly perturbed reduced order models perform better than Chebyshev Collocation approximation. At this point, it must be mentioned that a singular perturbation of the type (25) is also possible for Chebyshev Collocation approximation, but this involves considerable computational costs, as one does not have the analytical expressions for the approximate transfer functions obtained by using Chebyshev Collocation. Hence our claim that the approximate models proposed in this paper are better than the models obtained by using conventional schemes like Finite difference, Chebyshev Collocation, Chebyshev Galerkin and Chebyshev Tau is justified to a certain extent.

Having said the above, we list the advantages of using the models (10), (15), (25) and (26) proposed in this paper. First and the foremost, one can select two different model orders for approximating \(G_1(s)\) and \(G_2(s)\), which is not the case with any standard PDE solvers in numerical analysis. This is a big plus in terms of model reduction. The analytic expressions of the proposed models (10), (15), (25) and (26) are simple and easy to implement. This is not the case with Chebyshev Collocation, Chebyshev Galerkin and Chebyshev Tau, where it is difficult, if not impossible, to obtain the analytical expressions for the approximate transfer functions and setting up the differential equations for a numerical simulation in Chebyshev Collocation, Chebyshev Galerkin and Chebyshev Tau, could be a tedious task.

7 Appendix

We prove the expressions (20) and (21) in this appendix. The proof goes by mathematical induction. We use a new notation here. For a given \(n\) in (20) and (21), we refer to their partial fraction coefficients \(A_i\) as \(A^n_i\), \(B_i\) as \(B^n_i\) and \(K_1, K_2\) as \(K^n_1\) and \(K^n_2\) respectively.

We first prove (20). Let \(n = 2\), then we have

\[
G_2^2(s) = K_1^2 \frac{(s - z_1)}{(s - p_1)(s - p_2)}
\]  

(32)

where

\[
K_1^1 = -\frac{p_2 p_1}{z_1}
\]  

(33)
Figure 12: The approximate transfer function (10), plotted in solid line, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed dots.
Figure 13: The approximate transfer function (10), plotted in solid line, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed dots.
Figure 14: The approximate transfer function (10), plotted in solid line, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed dots.
Figure 15: The approximate transfer function (15), plotted in solid line, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed dots.
Figure 16: The approximate transfer function (15), plotted in solid line, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed dots.
Figure 17: The approximate transfer function (15), plotted in solid line, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed dots.
Figure 18: The Singularly perturbed model (25), plotted dashed dots, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed lines for a model of order $n = 3$. 
Figure 19: The Singularly perturbed model (25), plotted dashed dots, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed lines for a model of order $n = 5$. 
Figure 20: The Singularly perturbed model (25), plotted dashed dots, along with the true transfer function in the dotted line and Collocation approximation, refer to [10], in dashed lines for a model of order $n = 7$. 
By using partial fractions, (32) can be rewritten as

\[ G_1^1(s) = K_1^1 \left[ \frac{A_1^1}{s - p_1} + \frac{A_2^1}{s - p_2} \right] \]  

(34)

where

\[ A_1^1 = \frac{z_1 - p_1}{p_2 - p_1} \]  

(35)

\[ A_2^1 = \frac{z_1 - p_2}{p_1 - p_2} \]  

(36)

which is in agreement with (20) and (21).

Assume (20) to be true for \( n = 2, \ldots, N \), we claim it is true for \( n = N + 1 \), and thus for any \( n \) by induction.

Let \( n = N + 1 \). Then

\[ G_1^{N+1}(s) = K_1^{N+1} \frac{(s - z_1)(s - z_2) \cdots (s - z_N)}{(s - p_1)(s - p_2) \cdots (s - p_{N+1})} \]  

(37)

\[ = K_1^N \frac{(s - z_1)(s - z_2) \cdots (s - z_{N-1})}{(s - p_1)(s - p_2) \cdots (s - p_N)} \frac{s - z_N}{p_{N+1}} \]  

(38)

\[ = G_1^N \frac{p_{N+1}}{z_N} \frac{s - z_N}{s - p_{N+1}}. \]  

(39)

Since we have assumed (18) and (20) to be true for \( n = 2, \ldots, N \), using (20) in (39) we have

\[ G_1^{N+1}(s) = K_1^N \frac{p_{N+1}}{z_N} \left[ \sum_{i=1}^{N} \frac{A_i^N}{(s - p_i)} \right] \frac{s - z_N}{s - p_{N+1}} \]  

(40)

\[ = K_1^N \frac{p_{N+1}}{z_N} \left[ \sum_{i=1}^{N} \frac{A_i^N (s - z_N)}{(s - p_i)(s - p_{N+1})} \right]. \]  

(41)

As we did in (32) and (33), we rewrite (41) using partial fractions, which gives

\[ G_1^{N+1}(s) = K_1^N \frac{p_{N+1}}{z_N} \sum_{i=1}^{N} A_i^N \left[ \frac{(z_N - p_i)}{(p_{N+1} - p_i)} \frac{1}{(s - p_i)} \right. \]
\[ + \left. \frac{(z_N - p_{N+1})}{(p_i - p_{N+1})} \frac{1}{(s - p_{N+1})} \right]. \]  

(42)

which from (20) is

\[ G_1^{N+1}(s) = K_1^{N+1} \sum_{i=1}^{N} \frac{A_i^{N+1}}{(s - p_i)} + K_1^{N+1} \sum_{i=1}^{N} \frac{A_i^N (z_N - p_{N+1})}{(p_i - p_{N+1}) (s - p_{N+1})} \]  

(43)
Thus we are done, if we prove that
\[ \sum_{i=1}^{N} \frac{A_i^N(z_N - p_{N+1})}{(p_i - p_{N+1})} = A_{N+1}^N \]  \hspace{1cm} (44)

To prove (44), first note that, by equating (10) and (18), for any given \( n \), we have
\[ (s - z_1) \cdots (s - z_{m-1}) = \sum_{i=1}^{n} A_i^n(s - p_1) \cdots (s - p_{i-1})(s - p_{i+1}) \cdots (s - p_n). \]  \hspace{1cm} (45)

Hence by equating the coefficients of \( s^{n-1} \) in (45) we have
\[ \sum_{i=1}^{n} A_i^n = 1, \]  \hspace{1cm} (46)

for any given \( n \). Note that
\[ \frac{(z_N - p_{N+1})}{(p_i - p_{N+1})} = 1 - \frac{z_N - p_i}{p_{N+1} - p_i}. \]  \hspace{1cm} (47)

Hence using (47) in (44) we have
\[ \sum_{i=1}^{N} A_i^N \frac{(z_N - p_{N+1})}{(p_i - p_{N+1})} = \sum_{i=1}^{N} A_i^n - \sum_{i=1}^{N} A_i^n \frac{z_N - p_i}{p_{N+1} - p_i}. \]  \hspace{1cm} (48)

Note that
\[ A_i^N \frac{z_N - p_i}{p_{N+1} - p_i} = A_i^{N+1}, \]  \hspace{1cm} (49)

therefore using (46) and (49) in (48), we get
\[ \sum_{i=1}^{N} A_i^N \frac{(z_N - p_{N+1})}{(p_i - p_{N+1})} = 1 - \sum_{i=1}^{N} A_i^{N+1}, \]  \hspace{1cm} (50)

which by (46)
\[ \sum_{i=1}^{N} A_i^N \frac{(z_N - p_{N+1})}{(p_i - p_{N+1})} = 1 - (1 - A_{N+1}^{N+1}) \]
\[ = A_{N+1}^{N+1} \]  \hspace{1cm} (51)

Thus
\[ G_{1}^{N+1}(s) = K_{1}^{N+1} \sum_{i=1}^{N+1} A_i^{N+1} \frac{1}{(s - p_i)}. \]  \hspace{1cm} (53)

Hence (20) is proved.

A similar induction proof can be made for (21).
References


