Gramians of structured systems and an error bound for structure-preserving model reduction*

D.C. Sorensen and A.C. Antoulas†

e-mail:{sorensen@rice.edu, aca@rice.edu}

September 3, 2004

Abstract

In this paper a general framework is posed for defining the reachability and controllability gramians of structured linear dynamical systems. The novelty is that a formula for the gramian is given in the frequency domain. This formulation is surprisingly versatile and may be applied in a variety of structured problems. Moreover, this formulation enables a rather straightforward development of apriori error bounds for model reduction in the $\mathcal{H}_2$ norm. The bound applies to a reduced model derived from projection onto the dominant eigenspace of the appropriate gramian. The reduced models are structure preserving because they arise as direct reduction of the original system in the reduced basis. A derivation of the bound is presented and verified computationally on a second order system arising from structural analysis.

1 Introduction

The notion of reachability and observability gramians is well established in the theory of linear time invariant first order systems. However, there are several competing definitions of these quantities for higher order or structured systems. In particular, for second order systems, at least two different concepts have been proposed (see [6, 7]).

One of the main interests in defining these gramians is to develop a notion that will be suitable for model reduction via projection onto dominant invariant subspaces of the gramians. The goal is to provide model reductions that posses apriori error bounds analogous to those for balanced truncation of first order systems. The gramian definitions proposed in [6] for second order systems attempt to achieve a balanced reduction that preserves the second order structure of the system. The work reported in [7] and [8] is also concerned with preservation of second order structure. While the definitions in these investigations are reasonable and reduction schemes based upon the proposed gramians have been implemented, none of them have provided the desired error bounds.

In this paper, a fairly standard notion of gramian is proposed. The novelty is that a formula for the gramian is posed in the frequency domain. This formulation is surprisingly versatile and may be applied in a variety of structured problems. Moreover, this formulation in the frequency domain leads to apriori error bounds in the $\mathcal{H}_2$ norm in a rather straightforward way.

In the remainder of this paper, we shall lay out the general framework and show how the formulation leads to natural gramian definitions for a variety of structured problems. We then give a general derivation of an $\mathcal{H}_2$ norm error bound for model reduction based upon projection onto the dominant invariant subspace of the appropriate gramian. An example of a structure preserving reduction of a

*This work was supported in part by the NSF through Grants DMS-9972591, CCR-9988393 and ACI-0082645.

†ECE and CAAM, Rice University, Houston, Texas 77251-1892.

second order system is provided to experimentally verify the validity of the bound. The numerical results indicate that the new bound is rather tight for this example.

2 A framework for formulation of structured system gramians

Given is a system $\mathbf{\Sigma}$ described by the usual equations $\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$, $\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$, where $\mathbf{u}$, $\mathbf{x}$, $\mathbf{y}$ are the input, state, output and

$$\mathbf{\Sigma} = \left(\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right) \in \mathbb{R}^{(n+p)\times(n+m)}. \tag{1}$$

We will assume that the system is stable, that is, $\mathbf{A}$ has eigenvalues in the left-half of the complex plane. The *reachability gramian* of $\mathbf{\Sigma}$ is defined as

$$\mathcal{P} = \int_0^\infty \mathbf{x}(t)\mathbf{x}(t)^*\,dt, \tag{2}$$

where $\mathbf{x}$ is the solution of the state equation for $\mathbf{u}(t) = \delta(t)$. Using Parseval's theorem, the gramian can also be expressed in the frequency domain as

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^\infty \hat{\mathbf{x}}(i\omega)\hat{\mathbf{x}}^*(-i\omega)\,d\omega, \tag{3}$$

where $\hat{\mathbf{x}}$ denotes the transform of the time signal $\mathbf{x}$. Since the state due to an impulse is $\mathbf{x}(t) = e^{\mathbf{A}t}\mathbf{B}$ and equivalently $\hat{\mathbf{x}}(i\omega) = (i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}$, the gramian of $\mathbf{\Sigma}$ in time and in frequency is:

$$\mathcal{P} = \int_0^\infty e^{\mathbf{A}t}\mathbf{B}\mathbf{B}^* e^{\mathbf{A}^*t}\,dt = \frac{1}{2\pi}\int_{-\infty}^\infty (i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\,\mathbf{B}^*(-i\omega\mathbf{I} - \mathbf{A}^*)^{-1}\,d\omega. \tag{4}$$

This gramian has the following variational interpretation. Let $\mathbf{J}(\mathbf{u}, t_1, t_2) = \int_{t_1}^{t_2} \mathbf{u}^*(t)\mathbf{u}(t)\,dt$, i.e. $\mathbf{J}$ is the norm of the input function $\mathbf{u}$ in the time interval $[t_1,\ t_2]$. The following statements holds

$$\min_{\mathbf{u}} \mathbf{J}(\mathbf{u}, -\infty, 0) = \mathbf{x}_0^*\mathcal{P}^{-1}\mathbf{x}_0 \ \text{ subject to } \ \dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u},\ \mathbf{x}(0) = \mathbf{x}_0;$$

That is, the minimal energy required to steer the system from rest at $t = -\infty$, to $\mathbf{x}_0$ at time $t = 0$ is given by $\mathbf{x}_0^*\mathcal{P}^{-1}\mathbf{x}_0$.

By duality, we also define the *observability gramian* as follows:

$$\mathcal{Q} = \int_0^\infty e^{\mathbf{A}^*t}\mathbf{C}^*\mathbf{C}e^{\mathbf{A}t}\,dt = \frac{1}{2\pi}\int_{-\infty}^\infty (-i\omega\mathbf{I} - \mathbf{A}^*)^{-1}\mathbf{C}^*\,\mathbf{C}(i\omega\mathbf{I} - \mathbf{A})^{-1}\,d\omega.$$

2.1 Gramians for structured systems

We will now turn our attention to the following types of structured systems, namely: weighted, second-order, closed-loop and unstable systems. In terms of their transfer functions, these systems are as follows.

Weighted systems:	$\mathbf{G}_\mathbf{W}(s) = \mathbf{W}_o(s)\mathbf{G}(s)\mathbf{W}_i(s)$
Second order systems:	$\mathbf{G}_\mathbf{2}(s) = (s\mathbf{C}_1 + \mathbf{C}_0)(s^2\mathbf{M} + s\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}$
Systems in closed loop:	$\mathbf{G}_{\rm cl}(s) = \mathbf{G}(s)(\mathbf{I} + \mathbf{K}(s)\mathbf{G}(s))^{-1}$
Unstable systems:	$\mathbf{G}(s)$ with poles in $\mathbb{C}_+$.

2.2 Gramians for structured systems in frequency

In analogy with the case above, the reachability gramian of these systems will be defined as $\int \mathbf{x}\mathbf{x}^*$. In the case of input weighted systems with weight $\mathbf{W}$, the state of the system is $\mathbf{x}_{\mathbf{W}}(i\omega) = (i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\mathbf{W}(i\omega)$. Similarly for systems in a closed loop the system state is $\mathbf{x}_{\mathrm{cl}}(i\omega) = (i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}(\mathbf{I} + \mathbf{K}(i\omega)\mathbf{G}(i\omega))^{-1}$. In the case of second-order systems where $\mathbf{x}$ is position and $\dot{\mathbf{x}}$ the velocity, we can define two gramians, namely the position and velocity reachability gramians. Let the system in this case be described as follows:

$$\mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{y}(t) = \mathbf{C}_0\mathbf{x}(t) + \mathbf{C}_1\dot{\mathbf{x}}(t).$$

In this case we can define the position gramian

$$\mathcal{P}_0 = \int_0^\infty \mathbf{x}(t)\mathbf{x}^*(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\mathbf{x}}(i\omega)\hat{\mathbf{x}}^*(-i\omega)\,d\omega,$$

and the velocity gramian $\mathcal{P}_1 = \int \dot{\mathbf{x}}\dot{\mathbf{x}}^*$

$$\mathcal{P}_1 = \int_0^\infty \dot{\mathbf{x}}(t)\dot{\mathbf{x}}^*(t)\,dt = \frac{1}{2\pi}\int_{-\infty}^{\infty} (i\omega)\hat{\mathbf{x}}(i\omega)\hat{\mathbf{x}}^*(-i\omega)(-i\omega)\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \omega^2\hat{\mathbf{x}}(i\omega)\hat{\mathbf{x}}^*(-i\omega)\,d\omega.$$

$$\min_{\dot{\mathbf{x}}_0}\min_{\mathbf{u}} \mathbf{J}(\mathbf{u}, -\infty, 0) \text{ subject to } \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \mathbf{x}(0) = \mathbf{x}_0 \;\Rightarrow\; \mathbf{J}_{\min} = \mathbf{x}_0^*\mathcal{P}_0^{-1}\mathbf{x}_0,$$

$$\min_{\mathbf{x}_0}\min_{\mathbf{u}} \mathbf{J}(\mathbf{u}, -\infty, 0) \text{ subject to } \mathbf{M}\ddot{\mathbf{x}}(t) + \mathbf{D}\dot{\mathbf{x}}(t) + \mathbf{K}\mathbf{x}(t) = \mathbf{B}\mathbf{u}(t), \quad \dot{\mathbf{x}}(0) = \dot{\mathbf{x}}_0. \;\Rightarrow\; \mathbf{J}_{\min} = \dot{\mathbf{x}}_0^*\mathcal{P}_1^{-1}\dot{\mathbf{x}}_0,$$

Finally, for systems which are unstable (i.e. their poles are both in the right- and the left-half of the complex plane), the gramian is the following expression in the frequency domain

$$\mathcal{P}_{\mathrm{unst}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\mathbf{x}}(i\omega)\hat{\mathbf{x}}^*(-i\omega)\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} (i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\,\mathbf{B}^*(-i\omega\mathbf{I} - \mathbf{A}^*)^{-1}\,d\omega.$$

These gramians are summarized in table 1.

$\mathcal{P}_{\mathbf{W}} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\,\mathbf{W}(i\omega)\mathbf{W}^*(-i\omega)\,\mathbf{B}^*(-i\omega\mathbf{I} - \mathbf{A}^*)^{-1}\,d\omega$
$\mathcal{P}_{\mathbf{2}} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(-\omega^2\mathbf{M} + i\omega\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}\,\mathbf{B}^*(-\omega^2\mathbf{M}^* - i\omega\mathbf{D}^* + \mathbf{K}^*)^{-1}\,d\omega$
$\mathcal{P}_{\mathbf{cl}} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\,(\mathbf{I} + \mathbf{K}(i\omega)\mathbf{G}(i\omega))^{-1} \cdot \cdot\,(\mathbf{I} + \mathbf{K}^*(-i\omega)\mathbf{G}^*(-i\omega))^{-1}\,\mathbf{B}^*(-i\omega\mathbf{I} - \mathbf{A}^*)^{-1}\,d\omega$
$\mathcal{P}_{\mathbf{unst}} = \frac{1}{2\pi}\int_{-\infty}^{\infty}(i\omega\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}\,\mathbf{B}^*(-i\omega\mathbf{I} - \mathbf{A}^*)^{-1}\,d\omega$

Table 1: Gramians of structured systems

2.3 Gramians in the time domain

Our next goal is to express these gramians in the time domain as (part of the) solutions of appropriately defined Lyapunov equations. Recall that if $\mathbf{A}$ has eigenvalues in $\mathbb{C}_-$, the reachability gramian defined by (4) satisfies the following *Lyapunov equation*

$$\mathbf{A}\mathcal{P}(\mathbf{A},\mathbf{B}) + \mathcal{P}(\mathbf{A},\mathbf{B})\mathbf{A}^* + \mathbf{B}\mathbf{B}^* = \mathbf{0} \tag{5}$$

where for clarity the dependence of the gramian on $\mathbf{A}$ and $\mathbf{B}$ is shown explicitly. With this notation, given that the transfer function of the original system is $\mathbf{G}$, let the transfer function the weighted system be $\mathbf{W}_o\mathbf{G}\mathbf{W}_i$, where $\mathbf{W}_o$, $\mathbf{W}_i$ are the input and output weights respectively. The transfer function of the second-order system is $\mathbf{G_2}(s) = (\mathbf{C}_0 + \mathbf{C}_1 s)(\mathbf{M}s^2 + \mathbf{G}s + \mathbf{K})^{-1}\mathbf{B}$, while that of the closed loop system $\mathbf{G}_{\text{cl}} = \mathbf{G}(\mathbf{I} + \mathbf{K}\mathbf{G})^{-1}$. Given the state space realizations for the three systems $\mathbf{\Sigma_W}$, $\mathbf{\Sigma_2}$, $\mathbf{\Sigma}_{\text{cl}}$, collectively denoted as $\left(\begin{array}{c|c} \mathbf{A}_t & \mathbf{B}_t \\ \hline \mathbf{C}_t & \mathbf{D}_t \end{array}\right)$, the gramians are as follows.

<table>
<tr><td>$$\mathbf{\Sigma_W} = \left[\begin{array}{ccc|c} \mathbf{A}_o & \mathbf{B}_o\mathbf{C} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{B}\mathbf{C}_i & \mathbf{B}\mathbf{D}_i \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_i & \mathbf{B}_i \\ \hline \mathbf{C}_o & \mathbf{D}_o\mathbf{C} & \mathbf{0} & \mathbf{0} \end{array}\right]$$</td><td>$$\Rightarrow \mathcal{P}_{\mathbf{W}} = \left[\begin{array}{ccc} \mathbf{0} & \mathbf{I} & \mathbf{0} \end{array}\right] \mathcal{P}(\mathbf{A}_t, \mathbf{B}_t) \left[\begin{array}{c} \mathbf{0} \\ \mathbf{I} \\ \mathbf{0} \end{array}\right], \mathcal{Q}_{\mathbf{W}} = [\mathcal{Q}(\mathbf{C}_t, \mathbf{A}_t)]_{11}$$</td></tr>
<tr><td>$$\mathbf{\Sigma_2} = \left[\begin{array}{cc|c} \mathbf{0} & \mathbf{I} & \mathbf{0} \\ -\mathbf{M}^{-1}\mathbf{K} & -\mathbf{M}^{-1}\mathbf{D} & \mathbf{B} \\ \hline \mathbf{C}_1 & \mathbf{C}_0 & \mathbf{0} \end{array}\right]$$</td><td>$$\Rightarrow \mathcal{P}_0 = \left[\begin{array}{cc} \mathbf{I} & \mathbf{0} \end{array}\right] \mathcal{P}(\mathbf{A}_t, \mathbf{B}_t) \left[\begin{array}{c} \mathbf{I} \\ \mathbf{0} \end{array}\right], \mathcal{Q}_0 = [\mathcal{Q}(\mathbf{C}_t, \mathbf{A}_t)]_{11}$$
$$\Rightarrow \mathcal{P}_1 = \left[\begin{array}{cc} \mathbf{0} & \mathbf{I} \end{array}\right] \mathcal{P}(\mathbf{A}_t, \mathbf{B}_t) \left[\begin{array}{c} \mathbf{0} \\ \mathbf{I} \end{array}\right], \mathcal{Q}_1 = [\mathcal{Q}(\mathbf{C}_t, \mathbf{A}_t)]_{22}$$</td></tr>
<tr><td>$$\mathbf{\Sigma}_{\text{cl}} = \left[\begin{array}{cc|c} \mathbf{A} & -\mathbf{B}\mathbf{C}_c & \mathbf{B} \\ \mathbf{B}_c\mathbf{C} & \mathbf{A}_c & \mathbf{0} \\ \hline \mathbf{C} & \mathbf{0} & \mathbf{0} \end{array}\right]$$</td><td>$$\Rightarrow \mathcal{P}_{\text{cl}} = \left[\begin{array}{cc} \mathbf{I} & \mathbf{0} \end{array}\right] \mathcal{P}(\mathbf{A}_t, \mathbf{B}_t) \left[\begin{array}{c} \mathbf{I} \\ \mathbf{0} \end{array}\right], \mathcal{Q}_{\text{cl}} = [\mathcal{Q}(\mathbf{C}_t, \mathbf{A}_t)]_{11}$$</td></tr>
</table>

Lyapunov equations for unstable systems. The gramian defined above for unstable systems satisfies a Lyapunov equation as well. It is as follows:

$$\mathbf{A}\mathcal{P} + \mathcal{P}\mathbf{A}^* = \Pi\mathbf{Q}\Pi - (\mathbf{I} - \Pi)\mathbf{Q}(\mathbf{I} - \Pi)$$

where Π is the projection onto the **stable** eigenspace of $\mathbf{A}$. It turns out that $\Pi = \frac{1}{2}\mathbf{I} + \mathbf{S}$, where

$$\mathbf{S} = \frac{1}{2\pi} \int_{-\infty}^{\infty} (i\omega\mathbf{I} - \mathbf{A})^{-1}\, d\omega = \frac{i}{2\pi} \ln\left[(i\omega\mathbf{I} - \mathbf{A})(-i\omega\mathbf{I} - \mathbf{A})\right]\Big|_{\omega=\infty}.$$

3 A bound for the approximation error of structured systems

In order to introduce the class of systems under consideration we need the following notation. Let $\mathbf{Q}(s)$, $\mathbf{P}(s)$ be a polynomial matrices in s:

$$\mathbf{Q}(s) = \sum_{j=1}^{r} \mathbf{Q}_j s^j,\ \mathbf{Q}_j \in \mathbb{R}^{n\times n},\ \ \mathbf{P}(s) = \sum_{j=1}^{r-1} \mathbf{P}_j s^j,\ \mathbf{P}_j \in \mathbb{R}^{n\times m},$$

where $\mathbf{Q}$ is invertible and $\mathbf{Q}^{-1}\mathbf{P}$ is a strictly proper rational matrix. We will denote by $\mathbf{Q}(\frac{d}{dt})$, $\mathbf{P}(\frac{d}{dt})$ the differential operators

$$\mathbf{Q}(\frac{d}{dt}) = \sum_{j=1}^{r} \mathbf{Q}_j \frac{d^j}{dt^j},\ \ \mathbf{P}(\frac{d}{dt}) = \sum_{j=1}^{r-1} \mathbf{P}_j \frac{d^j}{dt^j}.$$

The systems are now defined by the following equations:

$$\mathbf{\Sigma}: \left\{ \begin{array}{rcl} \mathbf{Q}(\frac{d}{dt})\mathbf{x} & = & \mathbf{P}(\frac{d}{dt})\mathbf{u} \\ \mathbf{y}(t) & = & \mathbf{C}\mathbf{x}(t) \end{array} \right. \tag{6}$$

where $\mathbf{C} \in \mathbb{R}^{p\times n}$.

Here, we give a direct reduction of the above system based upon the dominant eigenspace of a gramian $\mathcal{P}$ that leads to an error bound in the $\mathcal{H}_2$ norm. An orthogonal basis for the dominant eigenspace of dimension k is used to construct a reduced model:

$$\hat{\mathbf{\Sigma}} : \begin{cases} \hat{\mathbf{Q}}(\frac{d}{dt})\hat{\mathbf{x}}(t) &= \hat{\mathbf{P}}(\frac{d}{dt})\mathbf{u}(t), \\ \hat{\mathbf{y}}(t) &= \hat{\mathbf{C}}\hat{\mathbf{x}}(t), \end{cases} \tag{7}$$

The *gramian* is defined as the gramian of $\mathbf{x}(t)$ when the input is an impulse:

$$\mathcal{P} := \int_0^\infty \mathbf{x}(t)\mathbf{x}(t)^* \mathrm{d}t.$$

Let

$$\mathcal{P} = \mathbf{V}\mathbf{S}\mathbf{V}^* \quad \text{with} \quad \mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2] \quad \text{and} \quad \mathbf{S} = diag(\mathbf{S}_1, \mathbf{S}_2),$$

be the eigensystem of $\mathcal{P}$, where the diagonal elements of $\mathbf{S}$ are in decreasing order, and $\mathbf{V}$ is orthogonal. The reduced model is derived from

$$\hat{\mathbf{Q}}_j = \mathbf{V}_1^*\mathbf{Q}_j\mathbf{V}_1, \quad \hat{\mathbf{P}}_j = \mathbf{V}_1^*\mathbf{P}_j, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{V}_1. \tag{8}$$

Our main result is the following:

Theorem 3.1 *The reduced model* $\hat{\mathbf{\Sigma}}$ *derived from the dominant eigenspace of the gramian* $\mathcal{P}$ *for* $\mathbf{\Sigma}$ *as described above satisfies*

$$\boxed{\|\mathbf{\Sigma} - \hat{\mathbf{\Sigma}}\|_{\mathcal{H}_2}^2 \leq \mathrm{trace}\,\{\mathbf{C}_2\mathbf{S}_2\mathbf{C}_2^*\} + \kappa\,\mathrm{trace}\,\{\mathbf{S}_2\}}$$

where κ *is a modest constant depending on* $\mathbf{\Sigma}$, $\hat{\mathbf{\Sigma}}$, *and the diagonal elements of* $\mathbf{S}_2$ *are the neglected smallest eigenvalues of* $\mathcal{P}$.

The following discussion will establish this result.

3.1 Details

It is readily verified that the *transfer function* for (6) in the frequency domain is

$$\mathbf{H}(s) = \mathbf{C}\mathbf{Q}^{-1}(s)\mathbf{P}(s)$$

Moreover, in the frequency domain, the input-to-$\mathbf{x}$ and the input-to-output maps are

$$\mathbf{x}(s) = \mathbf{Q}(s)^{-1}\mathbf{P}(s)\mathbf{u}(s), \quad \mathbf{y}(s) = \mathbf{H}(s)\mathbf{u}(s)$$

If the input is an impulse: $\mathbf{u}(t) = \delta(t)\mathbf{I}$ and $\mathbf{u}(s) = \mathbf{I}$,

$$\mathbf{x}(s) = \mathbf{Q}^{-1}(s)\mathbf{P}(s) \quad \text{and} \quad \mathbf{y}(s) = \mathbf{H}(s).$$

In the time domain

$$\int_0^\infty \mathbf{y}^*\mathbf{y}\mathrm{d}t = \mathrm{trace}\left\{\int_0^\infty \mathbf{y}\mathbf{y}^*\mathrm{d}t\right\} = \mathrm{trace}\left\{\int_0^\infty \mathbf{C}\mathbf{x}\mathbf{x}^*\mathbf{C}^*\mathrm{d}t\right\} = \mathbf{C}\mathcal{P}\mathbf{C}^*$$

Define $\mathbf{F}(s) := \mathbf{Q}^{-1}(s)\mathbf{P}(s)$. From the Parseval theorem, the above expression is equal to

$$\mathrm{trace}\left\{\int_0^\infty \mathbf{y}\mathbf{y}^*\mathrm{d}t\right\} = \mathrm{trace}\,\{\mathbf{C}\underbrace{\left(\frac{1}{2\pi}\int_{-\infty}^\infty \mathbf{F}(i\omega)\mathbf{F}(i\omega)^*\mathrm{d}\omega\right)\mathbf{C}^*}_{\mathcal{P}}\}.$$

Thus the gramian in the frequency domain is

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^\infty \mathbf{F}(i\omega)\mathbf{F}(i\omega)^*\mathrm{d}\omega.$$

Remark 3.1 Representation (6) is general as every system with strictly proper rational transfer function can be represented this way. Key for our considerations is the fact that the (square of the) $\mathcal{H}_2$ norm is given by $\mathbf{C}\mathcal{P}\mathbf{C}^*$. If instead of the output map, the input map is constant, the same framework can be applied by considering the transpose (dual) of the original system. ■

3.2 Reduction via the gramian

For model reduction, we consider again the eigen-decomposition of the symmetric positive definite matrix $\mathcal{P}$. Let

$$\mathcal{P} = \mathbf{V}\mathbf{S}\mathbf{V}^* \ \text{ with } \ \mathbf{V} = [\mathbf{V}_1, \ \ \mathbf{V}_2] \ \text{ and } \ \mathbf{S} = diag(\mathbf{S}_1, \ \mathbf{S}_2),$$

where the diagonal elements of $\mathbf{S}$ are in decreasing order, and $\mathbf{V}$ is orthogonal. The system is now transformed using $\mathbf{V}$ as in (8), to wit $\mathbf{Q}_j \leftarrow \mathbf{V}^*\mathbf{Q}_j\mathbf{V}$, $\mathbf{P}_j \leftarrow \mathbf{V}^*\mathbf{P}_j$, $\mathbf{C} \leftarrow \mathbf{C}\mathbf{V}$ which implies $\mathbf{F}(s) \leftarrow \mathbf{V}^*\mathbf{F}(s)$. In this new coordinate system the resulting gramian is diagonal. We now partition

$$\mathbf{Q}(s) = \begin{bmatrix} \mathbf{Q}_{11}(s) & \mathbf{Q}_{12}(s) \\ \mathbf{Q}_{21}(s) & \mathbf{Q}_{22}(s) \end{bmatrix}, \ \ [\mathbf{C}_1, \ \mathbf{C}_2] = \mathbf{C}\mathbf{V}, \ \ \mathbf{P}(s) = \begin{bmatrix} \mathbf{P}_1(s) \\ \mathbf{P}_2(s) \end{bmatrix} \ \text{ and } \ \mathbf{F}(s) = \begin{bmatrix} \mathbf{F}_1(s) \\ \mathbf{F}_2(s) \end{bmatrix}.$$

Note the relationship $\mathbf{Q}(s)\mathbf{F}(s) = \mathbf{P}(s)$. Let $\hat{\mathbf{Q}}(s) := \mathbf{Q}_{11}(s)$; since $\mathbf{S} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{F}(i\omega)\mathbf{F}(i\omega)^* \mathrm{d}\omega$, the following relationships hold

$$\begin{aligned} \mathbf{S}_1 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{F}_1(i\omega)\mathbf{F}_1(i\omega)^* \mathrm{d}\omega, \\ \mathbf{S}_2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{F}_2(i\omega)\mathbf{F}_2(i\omega)^* \mathrm{d}\omega, \\ \mathbf{0} &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathbf{F}_2(i\omega)\mathbf{F}_1(i\omega)^* \mathrm{d}\omega, \end{aligned}$$

while trace $\{\mathbf{S}_1\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\mathbf{F}_1(i\omega)\|_F^2 \mathrm{d}\omega$, and trace $\{\mathbf{S}_2\} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \|\mathbf{F}_2(i\omega)\|_F^2 \mathrm{d}\omega$. The reduced system is now constructed as follows:

$$\hat{\mathbf{Q}}_j = [\mathbf{Q}_j]_{11}, \ \hat{\mathbf{P}}_j = [\mathbf{P}_j]_{11}, \ \hat{\mathbf{C}} = \mathbf{C}_1.$$

Given $\hat{\mathbf{Q}}(s)$ as above we define $\hat{\mathbf{F}}$ by means of the equation $\hat{\mathbf{Q}}(s)\hat{\mathbf{F}}(s) = \mathbf{P}_1$. As a consequence of these definitions the gramian corresponding to the reduced system is

$$\hat{\mathcal{P}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{\mathbf{F}}(i\omega)\hat{\mathbf{F}}(i\omega)^* \mathrm{d}\omega,$$

and from the defining equation for $\mathbf{F}(s)$ we have

$$\mathbf{F}_1(s) = \mathbf{Q}_{11}(s)^{-1}[\mathbf{P}_1(s) - \mathbf{Q}_{12}(s)\mathbf{F}_2(s)] = \hat{\mathbf{F}}(s) - \mathbf{Q}_{11}(s)^{-1}\mathbf{Q}_{12}(s)\mathbf{F}_2(s).$$

Let $\mathbf{W}(s) := \mathbf{Q}_{11}(s)^{-1}\mathbf{Q}_{12}(s)$; if the reduced system has no poles on the imaginary axis, $\sup_\omega \|\mathbf{W}(i\omega)\|_2$ is finite. Thus,

$$\hat{\mathbf{F}}(s) = \mathbf{F}_1(s) + \mathbf{W}(s)\mathbf{F}_2(s).$$

3.3 Bounding the $\mathcal{H}_2$ norm of the error system

Applying the same input $\mathbf{u}$ to both the original and the reduced systems, let $\mathbf{y} = \mathbf{C}\mathbf{x}$, $\hat{\mathbf{y}} = \hat{\mathbf{C}}\hat{\mathbf{x}}$, be the resulting outputs. If we denote by $\mathbf{H}_e(s)$ the transfer function of the error system $\mathcal{E} = \Sigma - \hat{\Sigma}$, we have

$$\mathbf{y}(s) - \hat{\mathbf{y}}(s) = \mathbf{H}_e(s)\mathbf{u}(s) = \left[\mathbf{C}\mathbf{Q}(s)^{-1}\mathbf{P}(s) - \hat{\mathbf{C}}\hat{\mathbf{Q}}(s)^{-1}\hat{\mathbf{P}}(s)\right]\mathbf{u}(s).$$

The $\mathcal{H}_2$-norm in the of the error system is therefore

$$\begin{aligned}\|\mathcal{E}\|_{\mathcal{H}_2}^2 &= \text{trace}\,\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\mathbf{H}_e(i\omega)\mathbf{H}_e(i\omega)^*\mathrm{d}t\}\\ &= \underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty}\text{trace}\,\{\mathbf{C}\mathbf{F}(i\omega)(\mathbf{C}\mathbf{F}(i\omega))^*\}\mathrm{d}t}_{\eta_1} - 2\underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty}\text{trace}\,\{\mathbf{C}\mathbf{F}(i\omega)(\hat{\mathbf{C}}\hat{\mathbf{F}}(i\omega))^*\}\mathrm{d}t}_{\eta_2} + \\ &+ \underbrace{\frac{1}{2\pi}\int_{-\infty}^{\infty}\text{trace}\,\{\hat{\mathbf{C}}\hat{\mathbf{F}}(i\omega)(\hat{\mathbf{C}}\hat{\mathbf{F}}(i\omega))^*\}\mathrm{d}t}_{\eta_3}.\end{aligned}$$

Each of the three terms in this expression can be simplified as follows:

$$\eta_1 = \text{trace}\,\{\mathbf{C}_1 S_1 \mathbf{C}_1^*\} + \text{trace}\,\{\mathbf{C}_2 S_2 \mathbf{C}_2^*\},$$

$$\eta_2 = \text{trace}\,\{\mathbf{C}_1 S_1 \mathbf{C}_1^*\} + \frac{1}{2\pi}\int_{-\infty}^{\infty}\text{trace}\,\{\mathbf{C}\mathbf{F}(i\omega)\mathbf{F}_2(i\omega)^*\mathbf{W}(i\omega)^*\mathbf{C}_1^*\}\mathrm{d}t,$$

$$\begin{aligned}\eta_3 &= \text{trace}\,\{\mathbf{C}_1 S_1 \mathbf{C}_1^*\} + \frac{1}{2\pi}\int_{-\infty}^{\infty}2\text{trace}\,\{\mathbf{C}_1\mathbf{F}_1(i\omega)\mathbf{F}_2(i\omega)^*\mathbf{W}(i\omega)^*\mathbf{C}_1^*\}\mathrm{d}t + \\ &+ \frac{1}{2\pi}\int_{-\infty}^{\infty}\text{trace}\,\{(\mathbf{C}_1\mathbf{W}(i\omega)\mathbf{F}_2(i\omega))(\mathbf{C}_1\mathbf{W}(i\omega)\mathbf{F}_2(i\omega))^*\}\mathrm{d}t.\end{aligned}$$

Combining the above expressions we obtain

$$\|\mathcal{E}\|_{\mathcal{H}_2}^2 = \text{trace}\,\{\mathbf{C}_2 S_2 \mathbf{C}_2^*\} + \frac{1}{2\pi}\int_{-\infty}^{\infty}\text{trace}\,\{(\mathbf{C}_1\mathbf{W}(i\omega) - 2\mathbf{C}_2)\mathbf{F}_2(i\omega)(\mathbf{C}_1\mathbf{W}(i\omega)\mathbf{F}_2(i\omega))^*\}\mathrm{d}t.$$

The first term in the above expression is the $\mathcal{H}_2$ norm of the neglected term. The second term has the following upper bound

$$\sup_{\omega}\|(\mathbf{C}_1\mathbf{W}(i\omega))^*(\mathbf{C}_1\mathbf{W}(i\omega) - 2\mathbf{C}_2)\|_2\,\text{trace}\,\{\mathbf{S}_2\}.$$

This leads to the main result

$$\boxed{\|\mathcal{E}\|_{\mathcal{H}_2}^2 \leq \text{trace}\,\{\mathbf{C}_2\mathbf{S}_2\mathbf{C}_2^*\} + \kappa\,\text{trace}\,\{\mathbf{S}_2\} \quad \text{where} \quad \kappa = \sup_{\omega}\|(\mathbf{C}_1\mathbf{W}(i\omega))^*(\mathbf{C}_1\mathbf{W}(i\omega) - 2\mathbf{C}_2)\|_2} \qquad (9)$$

3.4 Special case: second-order systems

We shall now consider second-order systems. These are described by equations (7) where $\mathbf{Q}(s) = \mathbf{M}s^2 + \mathbf{D}s + \mathbf{K}$ and $\mathbf{P}(s) = \mathbf{B}$:

$$\mathbf{\Sigma}:\ \mathbf{M}\ddot{\mathbf{x}} + \mathbf{D}\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{B}\mathbf{u}, \quad \mathbf{y}(t) = \mathbf{C}\mathbf{x}(t), \qquad (10)$$

with $\mathbf{M}$, $\mathbf{D}$, $\mathbf{K} \in \mathbb{R}^{n\times n}$, $\mathbf{B} \in \mathbb{R}^{n\times m}$, $\mathbf{C} \in \mathbb{R}^{p\times n}$.

It is standard to convert this system to an equivalent first order linear time invariant (LTI) system and then to apply existing reduction techniques to reduce the first order system. A difficulty with this approach is that the second-order form is lost in the reduction process and there is a mixing of the state variables and their first derivatives. Several researchers (see e.g. [7], [8], [6]) have noted undesirable consequences and have endeavored to provide either direct reductions of the second-order form or structure preserving reductions of the equivalent first order system. This has required several alternative definitions of a gramian. However, while successful structure preserving reductions have been obtained, none of these possess apriori error bounds.

Here, we give a direct reduction of the second-order system based upon the dominant eigenspace of a gramian $\mathcal{P}$ that does lead to an error bound in the $\mathcal{H}_2$ norm. An orthogonal basis for the dominant eigenspace of dimension k is used to construct a reduced model in second-order form:

$$\hat{\boldsymbol{\Sigma}}:\ \hat{\mathbf{M}}\ddot{\hat{\mathbf{x}}}(t)+\hat{\mathbf{D}}\dot{\hat{\mathbf{x}}}(t)+\hat{\mathbf{K}}\hat{\mathbf{x}}(t)=\hat{\mathbf{B}}\mathbf{u}(t),\quad \hat{\mathbf{y}}(t)=\hat{\mathbf{C}}\hat{\mathbf{x}}(t).$$

The *gramian* is defined as before, i.e. $\mathcal{P}=\int_0^\infty \mathbf{x}\mathbf{x}^* dt$. Let $\mathcal{P}=\mathbf{V}\mathbf{S}\mathbf{V}^*$, with $\mathbf{V}=[\mathbf{V}_1,\mathbf{V}_2]$ and $\mathbf{S}=diag(\mathbf{S}_1,\mathbf{S}_2)$. The reduced model is derived by letting $\hat{\mathbf{M}}=\mathbf{V}_1^*\mathbf{M}\mathbf{V}_1$, $\hat{\mathbf{B}}=\mathbf{V}_1^*\mathbf{D}\mathbf{V}_1$, $\hat{\mathbf{K}}=\mathbf{V}_1^*\mathbf{K}\mathbf{V}_1$, $\hat{\mathbf{B}}=\mathbf{V}_1^*\mathbf{B}$, $\hat{\mathbf{C}}=\mathbf{C}\mathbf{V}_1$.

Remark 3.2 The above method applies equally to first-order systems, that is systems described by the equations $\dot{\mathbf{x}}=\mathbf{A}\mathbf{x}+\mathbf{B}\mathbf{u}$, $\mathbf{y}=\mathbf{C}\mathbf{x}+\mathbf{D}\mathbf{u}$. We will not pursue the details in this case here. ■

3.4.1 An illustrative example

The bound derived in the previous section involves the computation of the constant κ. The purpose of this section is to provide an example that will demonstrate that this constant is likely to be of reasonable magnitude. Our example is constructed to be representative of the structural analysis of a building under the assumption of proportional damping ($\mathbf{D}=\alpha\mathbf{M}+\beta\mathbf{K}$, for specified positive scalars α and β). In this case the matrices $\mathbf{M},\mathbf{D},\mathbf{K}$ may be simultaneously diagonalized. Moreover, since both $\mathbf{M}$ and $\mathbf{K}$ are positive definite, the system can be transformed to an equivalent one where $\mathbf{M}=\mathbf{I}$ and $\mathbf{K}$ is a diagonal matrix with positive diagonal entries.

The example may then be constructed by specification of the diagonal matrix $\mathbf{K}$, the proportionality constants α and β, and the vectors $\mathbf{B}$ and $\mathbf{C}$. We constructed $\mathbf{K}$ to have its smallest 200 eigenvalues specified as the smallest 200 eigenvalues of an actual building model of dimension $26,000$. These eigenvalues are in the range $[7.7, 5300]$. We augmented these with equally spaced eigenvalues $[5400 : 2000 : 400000]$ to obtain a diagonal matrix $\mathbf{K}$ of order $n=398$. We chose the proportionality constants $\alpha=.67$, $\beta=.0033$, to be consistent with the original building model. We specified $\mathbf{B}=\mathbf{C}^*$ to be vectors with all ones as entries. This is slightly inconsistent with the original building model but still representative. The eigenvalues of the second-order system resulting from this specification are shown in Figure 1

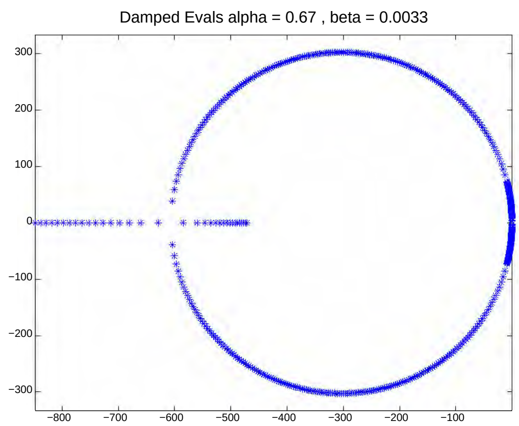

Figure 1: Eigenvalues of a proportionally damped structure

The gramian for proportional damping

To proceed we need to compute the gramian of this system. Recall from table 1 that

$$\mathcal{P} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (-\omega^2\mathbf{M} + i\omega\mathbf{D} + \mathbf{K})^{-1}\mathbf{B}\mathbf{B}^*(-\omega^2\mathbf{M}^* - i\omega\mathbf{D}^* + \mathbf{K}^*)^{-1}\, d\omega.$$

Since $\mathbf{M} = \mathbf{I}$, $\mathbf{D} = \text{diag}\,(d_1, \;\cdots\;, d_n)$ and $\mathbf{K} = \text{diag}\,(k_1, \;\cdots\;, k_n)$, the $(p,q)^{\text{th}}$ entry of the gramian is

$$\mathcal{P}_{pq} = \frac{1}{2\pi}\int_{-\infty}^{\infty} (-\omega^2 + i\omega d_p + k_p)^{-1} b_p b_q^* (-\omega^2 - i\omega d_q^* + k_q^*)^{-1}\, d\omega.$$

In order to compute this integral, we make use of the following partial fraction expansion:

$$\frac{b_p}{s^2 + d_p s + k_p} = \frac{\alpha_p}{s+\gamma_p} + \frac{\beta_p}{s+\delta_p}.$$

Then

$$\begin{aligned}\mathcal{P}_{pq} &= \int_0^{\infty} \left[\alpha_p e^{-\gamma_p t} + \beta_p e^{-\delta_p t}\right]\left[\alpha_q^* e^{-\gamma_q^* t} + \beta_q e^{-\delta_q^* t}\right] dt \\ &= \frac{\alpha_p\alpha_q^*}{\gamma_p+\gamma_q^*} + \frac{\alpha_p\beta_q^*}{\gamma_p+\delta_q^*} + \frac{\beta_p\alpha_q^*}{\delta_p+\gamma_q^*} + \frac{\beta_p\beta_q^*}{\delta_p+\delta_q^*}.\end{aligned}$$

With this formula, it is possible to explicitly construct the required gramian and diagonalize it. We set a tolerance of $\tau = 10^{-5}$, and truncated the second-order system to (a second-order system of) order k, such that $\sigma_{k+1}(\mathcal{P}) < \tau\cdot\sigma_1(\mathcal{P})$; the resulting reduced system has order $k = 51$.

$\|\mathbf{\Sigma}\|_{\mathcal{H}_2}^2$	3.9303e+000
$\|\hat{\mathbf{\Sigma}}\|_{\mathcal{H}_2}^2$	3.9302e+000
$\mathcal{H}_2$ norm of neglected system $\mathbf{C}_2\mathbf{S}_2\mathbf{C}_2^*$	4.3501e-005
κ	2.8725e+002
$\kappa\,\text{trace}\,(\mathbf{S}_2)$	1.7936e-003
Relative error bound	4.6743e-004
Computed relative error $\frac{\|\mathcal{E}\|_{\mathcal{H}_2}^2}{\|\mathbf{\Sigma}\|_{\mathcal{H}_2}^2}$	1.2196e-005

These results indicate that the constant κ in (9) is of moderate size and that the bound gives a reasonable error prediction.

A graphical illustration of the frequency response of the reduced model (order 51) compared to full system (order 398) is shown in Figure 2.

4 Summary

We have presented a unified way of defining gramians for structured systems, in particular, weighted, second-order, closed loop and unstable systems. The key is to start with the frequency domain. Consequently we examined the reduction of a high-order (structured) system based upon the dominant eigenspace of an appropriately defined gramian, that preserves the high-order form. An error bound in the $\mathcal{H}_2$ norm for this reduction was derived. An equivalent definition of the gramian was obtained through a Parseval relationship and this was key to the derivation of the bound. Here, we just sketched the derivations. Full details and computational issues will be reported in the future.

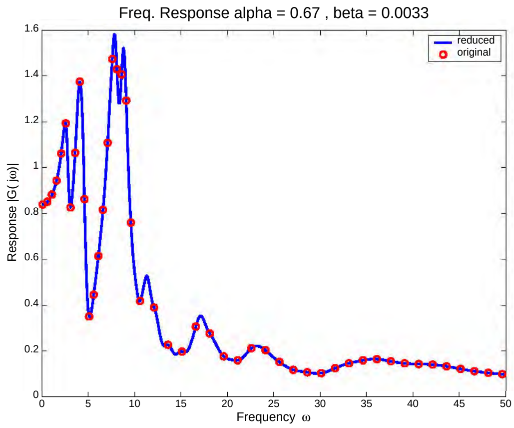

Figure 2: Frequency response of reduced model (order 51) compared to full system (order 398).

References

[1] A.C. Antoulas, *Approximation of large-scale dynamical systems*, SIAM Book series "Advances in Design and Control", Philadelphia (2004) (in press).

[2] A.C. Antoulas and D.C. Sorensen, *Lanczos, Lyapunov and Inertia*, Linear Algebra and Its Applications, **326**: 137-150 (2001).

[3] A.C. Antoulas, D.C. Sorensen, and S. Gugercin, *A survey of model reduction methods for large-scale systems*, Contemporary Mathematics, vol. 280, (2001), p. 193-219.

[4] S. Gugercin and A.C. Antoulas, *On balancing related model reduction methods and the corresponding error*, Int. Journal of Control, accepted for publication (2003).

[5] D.C. Sorensen and A.C. Antoulas, *The Sylvester equation and approximate balanced reduction*, Linear Algebra and Its Applications. Fourth Special Issue on Linear Systems and Control, Edited by V. Blondel, D. Hinrichsen, J. Rosenthal, and P.M. van Dooren., **351-352**: 671-700 (2002).

[6] D.G. Meyer and S. Srinivasan, *Balancing and model reduction for second-order form linear systems*, IEEE Trans. Automatic Control, **AC-41**: 1632-1644 (1996).

[7] Y. Chahlaoui, D. Lemonnier, A. Vandendorpe, and P. Van Dooren, *Second-order structure preserving model reduction*, Proc. MTNS, Leuven (2004).

[8] K. Chahlaoui, D. Lemonnier, K. Meerbergen, A. Vandendorpe, and P. Van Dooren, *Model reduction of second-order systems*, Proc. International Symposium Math. Theory. Netw. Syst., Paper 26984–4 (2002)