



Math 673

# Multigrid Methods: A Mostly Matrix-Based Approach

## Chapter 03: Fourier Analysis of the Two-Grid Algorithm

Abner J. Salgado and Steven M. Wise

asalgad1@utk.edu — swise1@utk.edu  
University of Tennessee

Fall 2024



## Chapter 03, Part 2 of 3

# Fourier Analysis of the Two-Grid Algorithm



# The Gauss-Seidel Smoother



## The Gauss-Seidel Smoother

Recall from our model problem that we want to solve the linear system

$$A_1 \mathbf{u}_1 = \mathbf{f}_1^\square \in \mathbb{R}^{n_1}. \quad (1)$$

To approximate the solution of Equation (1) we will next consider the forward Gauss-Seidel method. To define the method, we again split  $A_1$  into

$$A_1 = D - U - L. \quad (2)$$

where

$$D = \begin{bmatrix} \frac{2}{h_1} & & & & \\ & \frac{2}{h_1} & & & \\ & & \ddots & & \\ & & & \frac{2}{h_1} & \\ & & & & \frac{2}{h_1} \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}, \quad (3)$$

$$U = \begin{bmatrix} 0 & \frac{1}{h_1} & & & \\ & 0 & \frac{1}{h_1} & & \\ & & \ddots & \ddots & \\ & & & 0 & \frac{1}{h_1} \\ & & & & 0 \end{bmatrix} \in \mathbb{R}^{n_1 \times n_1}. \quad (4)$$



Of course,  $L = U^T$ . The Gauss-Seidel method can be expressed as

$$\mathbf{u}_1^{(\sigma+1)} = (D - L)^{-1} U \mathbf{u}_1^{(\sigma)} + (D - L)^{-1} \mathbf{f}_1^\square. \quad (5)$$

Equivalently,

$$\mathbf{u}_1^{(\sigma+1)} = \mathbf{u}_1^{(\sigma)} + (D - L)^{-1} \left( \mathbf{f}_1^\square - A_1 \mathbf{u}_1^{(\sigma)} \right). \quad (6)$$

In our two-grid terminology,

$$S_1 = (D - L)^{-1}$$

and

$$K_1 = I_1 - S_1 A_1 = I_1 - (D - L)^{-1} A_1.$$

Clearly  $S_1 \neq S_1^T$  and  $K_1 \neq K_1^*$ .



## The Gauss-Seidel Smoother

For the particular application, the Gauss-Seidel method may be written in component form as

$$u_{1,i}^{(\sigma+1)} = \frac{h_1}{2} \left\{ \frac{1}{h_1} u_{1,i-1}^{(\sigma+1)} + \frac{1}{h_1} u_{1,i+1}^{(\sigma)} \right\} + \frac{h_1}{2} f_{1,i}^{\square},$$

stepping through components from  $i = 1$ , in order, to  $i = n_1$ . We use component-wise updates immediately after they are generated. Therefore, the order in which we pass through the components is important.



## Theorem (Eigen-Pairs of Gauss-Seidel)

Let  $K_1 = I_1 - (D - L)^{-1} A_1$  be the error transfer matrix for the Gauss-Seidel method applied to the model problem (1). The eigenvectors of  $K_1$  are

$$\left[ \mathbf{w}_1^{(k)} \right]_i = w_{1,i}^{(k)} = [\cos(k\pi h_1)]^i \sin(k\pi x_{1,i}), \quad 1 \leq i \leq n_1,$$

for  $k = 1, \dots, n_1$ . The eigenvalues of  $K_1$  are

$$\nu_1^{(k)} = \cos^2(k\pi h_1), \quad 1 \leq k \leq n_1. \quad (7)$$

## Proof.

We begin by writing

$$K_1 = (D - L)^{-1} U.$$

Then

$$K_1 \mathbf{w}_1^{(k)} = \nu_1^{(k)} \mathbf{w}_1^{(k)}$$

if and only if

$$U \mathbf{w}_1^{(k)} = \nu_1^{(k)} (D - L) \mathbf{w}_1^{(k)}.$$

The rest of the details are left to the reader. □



## Remark

*Notice that the eigenvectors of the error transfer matrix for the forward Gauss-Seidel method are different from those of the stiffness matrix. This complicates the analysis of the multigrid method when Gauss-Seidel smoothing is used. We will not consider this smoother further until we develop new analysis techniques based on subspace decompositions. We will talk about this again in a later chapter.*





# The Smoothing Effect



Low frequency modes are, essentially, those that make sense on the coarse grid. Let us make explicit now the structure of the coarse grid.

### Definition (Nested and Uniform Grids)

We say that the **two-level grids are uniform and nested** iff  $n_1 + 1 \geq 4$ ;  $n_1 + 1$  is even; and

$$n_0 := \frac{n_1 + 1}{2} - 1.$$

In this case, we define the uniform grid sizes via

$$h_\ell = \frac{1}{n_\ell + 1}, \quad \ell = 0, 1,$$

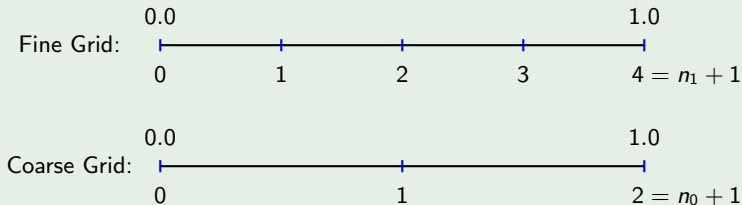
and the grid points via,

$$x_{\ell,i} = i \cdot h_\ell, \quad \ell = 0, 1.$$



## Example

For example, suppose  $n_1 = 3$ . A uniform and nested two-level grid is shown in the figure below.



**Figure:** Fine and coarse grids for  $n_1 = 3$ , where the two-level grids are uniform and nested.



A quick recall of the property of the stiffness matrix.

### Theorem (Stiffness Matrix is SPD)

*The level-1 stiffness matrix,  $A_1 \in \mathbb{R}^{n_1 \times n_1}$ , is SPD. Its eigenvalues are*

$$\lambda_1^{(k)} = \frac{4}{h_1} \sin^2 \left( \frac{k\pi h_1}{2} \right) = \frac{2}{h_1} (1 - \cos(k\pi h_1)). \quad (8)$$

*$k = 1, 2, \dots, n_1$ , and the corresponding eigenvectors are*

$$\left[ \mathbf{v}_1^{(k)} \right]_i = v_{1,i}^{(k)} = \sin(k\pi x_{1,i}), \quad 1 \leq i \leq n_1. \quad (9)$$



## Definition (High and Low Frequency Modes)

Suppose the two-level grids are uniform and nested. Consider an expansion of the form

$$\mathbf{v}_1 = \sum_{k=1}^{n_1} c_k \mathbf{v}_1^{(k)}, \quad (10)$$

where  $\mathbf{v}_1^{(k)}$  is the  $k^{\text{th}}$  eigenvector defined in (9). We say that the  $k^{\text{th}}$  mode,  $c_k \mathbf{v}_1^{(k)}$ , is of **high frequency** iff

$$n_0 + 1 = \frac{n_1 + 1}{2} \leq k \leq n_1.$$

Otherwise, we say that the mode is of **low frequency**.



Another quick recall.

### Theorem (Eigen-Pairs of Damped Jacobi)

*Let  $K_1 = I_1 - \omega D^{-1}A_1$  be the error transfer matrix for the damped Jacobi method applied to the model problem (1). The eigenvectors of  $K_1$  are the same as those for the level-1 stiffness matrix,  $A_1$ , that is,*

$$\left[ \mathbf{v}_1^{(k)} \right]_i = v_{1,i}^{(k)} = \sin(k\pi x_{1,i}), \quad 1 \leq i \leq n_1,$$

*for  $k = 1, \dots, n_1$ . The eigenvalues of  $K_1$  are*

$$\begin{aligned} \mu_1^{(k)}(\omega) &= \omega \cos(k\pi h_1) + 1 - \omega \\ &= 1 - 2\omega \sin^2\left(\frac{k\pi h_1}{2}\right), \quad 1 \leq k \leq n_1. \end{aligned} \quad (11)$$



## Theorem

Suppose that  $\mu_1^{(k)}(\omega)$  is the  $k^{\text{th}}$  eigenvalue (Equation (11)) of the error transfer matrix,  $K_1$ , of the damped Jacobi smoother and  $0 < \omega \leq 1$ . The quantity

$$S(\omega) = \max_{\frac{n_1+1}{2} \leq k \leq n_1} \left| \mu_1^{(k)}(\omega) \right|,$$

is minimized by

$$\omega = \omega_0 := \frac{2}{3},$$

in which case

$$\left| \mu_1^{(k)}(\omega_0) \right| \leq \frac{1}{3}, \quad (12)$$

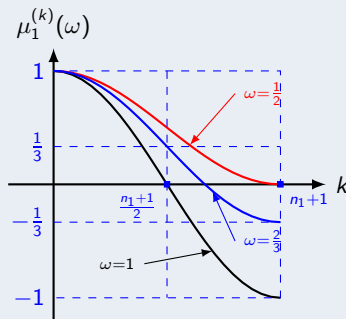
for all  $\frac{n_1+1}{2} \leq k \leq n_1$ . More generally, if  $0 < \omega \leq 1$ , then

$$\left| \mu_1^{(k)}(\omega) \right| < 1. \quad (13)$$



## Proof.

The proof follows by a careful examination of the plots of  $\mu_1^{(k)}(\omega) = 1 - 2\omega \sin^2\left(\frac{k\pi h_1}{2}\right)$  for various values of  $\omega$ . See the figure below.



**Figure:** Plots of the eigenvalues of  $K_1$ ,  $\mu_1^{(k)}(\omega) = 1 - 2\omega \sin^2\left(\frac{k\pi h_1}{2}\right)$ , as functions of  $k$ , for various values of  $\omega \in (0, 1]$ .







## Remark

*Recall, with  $\omega = \omega_0$ , we have*

$$\mathbf{e}_1^{(\sigma+1)} = \sum_{k=1}^{n_1} \mu_1^{(k)}(\omega_0) \epsilon_k^{(\sigma)} \mathbf{v}_1^{(k)}$$

*for the error after one smoothing iteration by the damped Jacobi method. High-frequency modes will be damped faster than those of low-frequency. In fact the modes  $\frac{n_1+1}{2} \leq k \leq n_1$  will be reduced by at least  $\frac{1}{3}$  after a single smoothing iteration.*

## The Two Multigrid Principles



### 1<sup>st</sup> MULTIGRID PRINCIPLE:

Many classical iterative methods have an error smoothing property – namely, high-frequency modes of the error are damped much more rapidly than those of low-frequency – but converge very slowly, especially as  $h_1 \rightarrow 0$ .

### 2<sup>nd</sup> MULTIGRID PRINCIPLE:

Low-frequency information is well approximated on a coarse grid.



## Remark

*We observed that a smoothed error is well-approximated on a coarse grid. By smoothed, we mean that the high frequency modes are greatly diminished. We can show that if the error is smooth, then the residual is almost as smooth. Let us explain.*

*Suppose that*

$$\mathbf{e}_1^{(\sigma)} = \sum_{k=1}^{n_1} \epsilon_k^{(\sigma)} \mathbf{v}_1^{(k)}.$$

*Since  $\mathbf{r}_1^{(\sigma)} = A_1 \mathbf{e}_1^{(\sigma)}$ ,*

$$\begin{aligned} \mathbf{r}_1^{(\sigma)} &= \sum_{k=1}^{n_1} \epsilon_k^{(\sigma)} A_1 \mathbf{v}_1^{(k)} \\ &= \sum_{k=1}^{n_1} \epsilon_k^{(\sigma)} \lambda_1^{(k)} \mathbf{v}_1^{(k)}. \end{aligned} \tag{14}$$

*If the high frequency modes of the error are totally absent, then, of course, they will be absent from the residual as well.*



## Remark (Cont.)

However, we note that as  $h_1 \rightarrow 0$ , the high frequency eigenvalues can become large. See the figure below. Thus high frequency components of the error can be amplified in the residual somewhat.

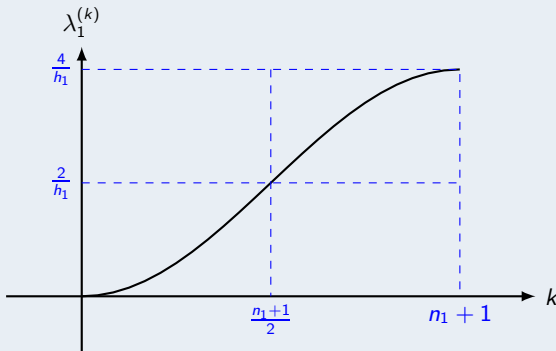


Figure: Eigenvalues of the level-1 stiffness matrix  $A_1$ .



# Prolongation and Restriction Operators



## Prolongation and Restriction Operators

We are going to approach the definition of the prolongation and the restriction matrices in what might seem like reverse order. We first define the prolongation matrix  $P_0$ , then we set  $R_0 = P_0^T$ .

Our construction comes from the FEM point of view. Let us define some objects associated to the coarse grid. Set

$$V_0 := \left\{ v \in C^0([0, 1]) \mid v(0) = v(1) = 0, v|_{T_{0,i}} \in \mathbb{P}_1(T_{0,i}), 1 \leq i \leq n_0 \right\}, \quad (15)$$

where the coarse grid is comprised of  $n_0$  equally sized intervals

$$T_{0,i} := (x_{0,i-1}, x_{0,i}), \quad i = 1, \dots, n_0,$$

and the coarse grid point set,  $\{x_{0,i}\}_{i=0}^{n_0+1}$ , is as defined via

$$x_{0,i} = i \cdot h_0, \quad h_0 := \frac{1}{n_0 + 1}.$$



Similar to level-1 hat functions, we can define level-0 hat functions.

### Definition (Hat Function)

For  $i = 1, \dots, n_0$ , define  $\psi_{0,i} \in V_0$  via

$$\psi_{0,i}(x_{0,j}) = \delta_{i,j}, \quad 1 \leq j \leq n_0.$$

$\psi_{0,i}$  is called a **level-0 hat function**.

The three hat functions for  $n_1 = 3$  are shown in the figure on the next slide.

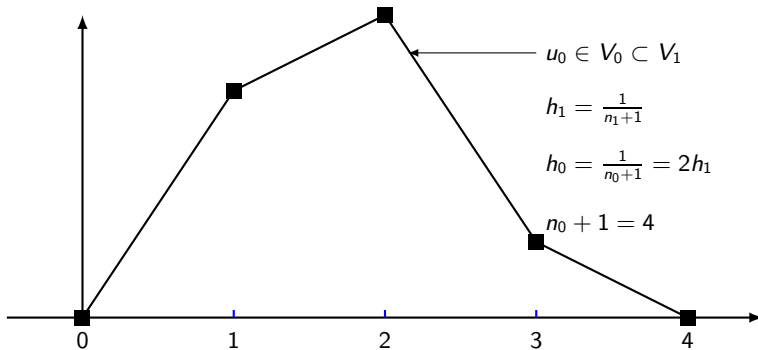


Figure: A piecewise linear function defined on the coarse grid, where  $n_0 = 3$ .



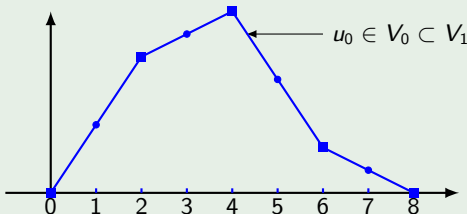


## Example

Now, suppose  $u_0 \in V_0$  is piecewise linear. See, for example, the figure below, where  $n_0 = 3$ . Recall, based on our uniform nested grids assumption

$$n_0 = \frac{n_1 + 1}{2} - 1 \Leftrightarrow n_1 = 2(n_0 + 1) - 1.$$

In the figure,  $n_1 = 7$ . Since  $V_0 \subset V_1$ ,  $u_0 \in V_1$ . This coarse grid function can be trivially represented on the fine grid, as seen in the figure below.



**Figure:** The coarse grid function from Figure 4 as it would be represented on the fine grid.



## Example (Cont.)

Our task is to find a matrix that maps the 3 coarse grid degrees of freedom (DOFs) into the 7 fine grid DOFs. The matrix  $P_0 \in \mathbb{R}^{n_1 \times n_0}$  will have the action

$$P_0 \mathbf{u}_0 = \mathbf{u}_1 \in \mathbb{R}^{n_1},$$

where  $\mathbf{u}_0 \in \mathbb{R}^{n_0}$  is coordinate representation of  $u_0 \in V_0$  in the level-0 hat function basis and  $\mathbf{u}_1 \in \mathbb{R}^{n_1}$  is the representation of  $u_0 \in V_1$  in the level-1 hat function basis. For this example ( $n_0 = 3, n_1 = 7$ ) the desired matrix is clearly

$$P_0 = \begin{bmatrix} \frac{1}{2} & & & & & & \\ \mathbf{1} & & & & & & \\ \frac{1}{2} & & & & & & \\ & \frac{1}{2} & & & & & \\ & \mathbf{1} & & & & & \\ & \frac{1}{2} & & & & & \\ & & \frac{1}{2} & & & & \\ & & \mathbf{1} & & & & \\ & & \frac{1}{2} & & & & \end{bmatrix} \in \mathbb{R}^{7 \times 3} \implies R_0 = \begin{bmatrix} \frac{1}{2} & & & & & & \\ & \mathbf{1} & & & & & \\ & & \frac{1}{2} & & & & \\ & & & \mathbf{1} & & & \\ & & & & \frac{1}{2} & & \\ & & & & \frac{1}{2} & & \\ & & & & & \mathbf{1} & \\ & & & & & & \frac{1}{2} \end{bmatrix} \in \mathbb{R}^{3 \times 7}.$$

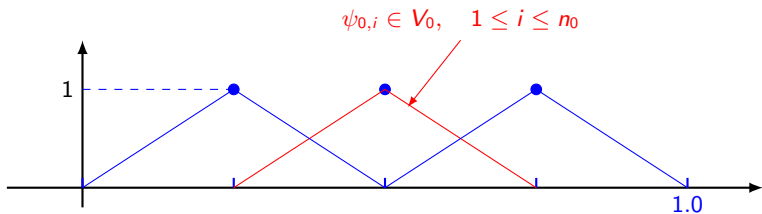


Figure: Level-0 hat function basis.

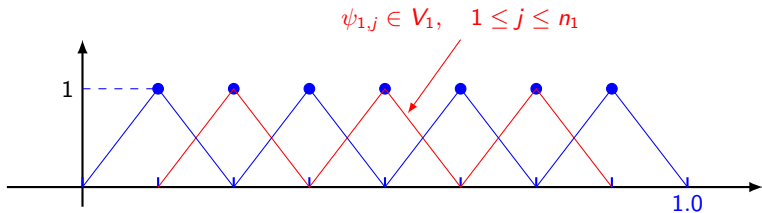


Figure: Level-1 hat function basis.



## Definition (Prolongation and Restriction)

Suppose that the positive integers  $n_0$  and  $n_1$  satisfy

$$n_1 = 2(n_0 + 1) - 1.$$

The action of  $P_0 \in \mathbb{R}^{n_1 \times n_0}$  on the arbitrary vector  $\mathbf{u}_0 \in \mathbb{R}^{n_0}$  is defined as follows

$$\begin{aligned} [P_0 \mathbf{u}_0]_1 &: = \frac{1}{2} u_{0,1}, \\ [P_0 \mathbf{u}_0]_{n_1} &: = \frac{1}{2} u_{0,n_0}, \\ [P_0 \mathbf{u}_0]_{2i} &: = u_{0,i}, \quad 1 \leq i \leq n_0, \\ [P_0 \mathbf{u}_0]_{2i+1} &: = \frac{1}{2} (u_{0,i} + u_{0,i+1}), \quad 1 \leq i \leq n_0 - 1. \end{aligned}$$

We define

$$R_0 = P_0^T \in \mathbb{R}^{n_0 \times n_1}. \quad (16)$$



## Theorem (Characterization of Prolongation Components)

*Suppose that the positive integers  $n_0$  and  $n_1$  satisfy*

$$n_1 = 2(n_0 + 1) - 1$$

*and the prolongation operator,  $P_0 = [p_{0,i,j}] \in \mathbb{R}^{n_1 \times n_0}$ , is as in Definition 11.*

*Suppose that*

$$B_\ell := \{\psi_{\ell,j}\}_{j=1}^{n_\ell}, \quad \ell = 0, 1.$$

*are the hat function bases of  $V_\ell$ ,  $\ell = 0, 1$ , respectively. Then the numbers  $p_{0,i,j}$ ,  $1 \leq i \leq n_1$ ,  $1 \leq j \leq n_0$  are the unique numbers satisfying*

$$\psi_{0,j} = \sum_{i=1}^{n_1} p_{0,i,j} \psi_{1,i}. \quad (17)$$



## Theorem (Characterization of Prolongation Components Cont.)

*Thus, if  $u_0 \in V_0 \subset V_1$  and*

$$u_0 = \sum_{j=1}^{n_0} u_{0,j} \psi_{0,j},$$

*then, the unique representation of  $u_0$  in the basis  $B_1$  is given by*

$$\begin{aligned} u_0 &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_0} p_{0,i,j} u_{0,j} \psi_{1,i} \\ &= \sum_{i=1}^{n_1} [P_0 u_0]_i \psi_{1,i}. \end{aligned} \tag{18}$$



## Proof.

Since  $V_0 \subset V_1$ , each level-0 hat basis function satisfies  $\psi_{0,j} \in V_1$ . Therefore, since  $B_1 = \{\psi_{1,i}\}_{i=1}^{n_1}$  is a basis for  $V_1$ , there exist unique numbers, which we conveniently label  $p_{0,i,j}$ ,  $1 \leq i \leq n_1$ ,  $1 \leq j \leq n_0$ , such that

$$\psi_{0,j} = \sum_{i=1}^{n_1} p_{0,i,j} \psi_{1,i}, \quad \forall j = 1, \dots, n_0.$$

We leave it as an exercise for the reader to prove that these are exactly the elements of the matrix  $P_0$ . This proves Equation (17).

Now, suppose that  $u_0 \in V_0 \subset V_1$  is arbitrary and

$$u_0 = \sum_{j=1}^{n_0} u_{0,j} \psi_{0,j}.$$



## Proof (Cont.)

The coordinate vector of this function in the hat-function basis  $B_0 = \{\psi_{0,j}\}_{j=1}^{n_0}$  is precisely

$$\mathbf{u}_0 = \begin{bmatrix} u_{0,1} \\ u_{0,2} \\ \vdots \\ u_{0,n_0} \end{bmatrix}.$$





## Proof (Cont.)

Using (17), we have

$$\begin{aligned} u_0 &= \sum_{j=1}^{n_0} u_{0,j} \psi_{0,j} \\ &= \sum_{j=1}^{n_0} u_{0,j} \sum_{i=1}^{n_0} u_{0,j} \psi_{0,j} \\ &= \sum_{i=1}^{n_1} \sum_{j=1}^{n_0} p_{0,i,j} u_{0,j} \psi_{1,i} \\ &= \sum_{i=1}^{n_1} [P_0 u_0]_i \psi_{1,i}. \end{aligned}$$

The result is proven. □



## Remark

*Suppose that  $\mathbf{u}_0 \in \mathbb{R}^{n_0}$  is the unique coordinate vector of the function  $u_0 \in V_0$  with respect to the basis  $B_0$ . Then  $P_0 \mathbf{u}_0$  is the unique coordinate vector of the function  $u_0$  with respect to the basis  $B_1$ .*



## Theorem (Verification of the Galerkin Condition)

Suppose the two-level grids are uniform and nested. Let  $P_0, R_0$  be defined as above, with the level-1 stiffness matrix  $A_1$  defined by

$$a_{1,i,j} = (\psi'_{1,j}, \psi'_{1,i})_{L^2(0,1)} = (\psi'_{1,i}, \psi'_{1,j})_{L^2(0,1)}. \quad (19)$$

The matrix  $A_0 = [a_{0,i,j}] \in \mathbb{R}^{n_0 \times n_0}$  satisfies the Galerkin condition, that is,

$$A_0 = R_0 A_1 P_0,$$

iff

$$a_{0,i,j} = (\psi'_{0,j}, \psi'_{0,i})_{L^2(0,1)} = (\psi'_{0,i}, \psi'_{0,j})_{L^2(0,1)}.$$

Either way,

$$A_0 = \begin{bmatrix} \frac{2}{h_0} & -\frac{1}{h_0} & & & \\ -\frac{1}{h_0} & \frac{2}{h_0} & -\frac{1}{h_0} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{h_0} & \frac{2}{h_0} & -\frac{1}{h_0} \\ & & & -\frac{1}{h_0} & \frac{2}{h_0} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0}.$$



## Proof.

We prove only one direction. The other will be obvious.

( $\Leftarrow$ ): Recall that, from (19), the level-1 stiffness matrix is defined via

$$[A_1]_{i,j} := (\psi'_{1,j}, \psi'_{1,i})_{L^2(0,1)}.$$

Defining the level-0 stiffness matrix analogously and using (17), we have

$$\begin{aligned} [A_0]_{i,j} &= (\psi'_{0,j}, \psi'_{0,i})_{L^2(0,1)} \\ &= (\psi'_{0,i}, \psi'_{0,j})_{L^2(0,1)} \\ &\stackrel{(17)}{=} \left( \sum_{k=1}^{n_1} p_{0,k,i} \psi'_{1,k}, \sum_{\ell=1}^{n_1} p_{0,\ell,j} \psi'_{1,\ell} \right)_{L^2(0,1)} \\ &= \sum_{k=1}^{n_1} \sum_{\ell=1}^{n_1} p_{0,k,i} (\psi'_{1,k}, \psi'_{1,\ell})_{L^2(0,1)} p_{0,\ell,j} \\ &= \sum_{k=1}^{n_1} \sum_{\ell=1}^{n_1} [R_0]_{i,k} [A_1]_{k,\ell} [P_0]_{\ell,j} \\ &= [R_0 A_1 P_0]_{i,j}. \end{aligned} \tag{20}$$



## Proof (Cont.)

Finally, since

$$a_{0,i,j} = (\psi'_{0,j}, \psi'_{0,i})_{L^2(0,1)} = (\psi'_{0,i}, \psi'_{0,j})_{L^2(0,1)},$$

it is an easy exercise to show that

$$A_0 = \begin{bmatrix} \frac{2}{h_0} & -\frac{1}{h_0} & & & \\ -\frac{1}{h_0} & \frac{2}{h_0} & -\frac{1}{h_0} & & \\ & \ddots & \ddots & \ddots & \\ & & -\frac{1}{h_0} & \frac{2}{h_0} & -\frac{1}{h_0} \\ & & & -\frac{1}{h_0} & \frac{2}{h_0} \end{bmatrix} \in \mathbb{R}^{n_0 \times n_0},$$

using the same techniques used to construct the level-1 stiffness matrix. □



## Corollary

*Suppose the two-level grids are uniform and nested and the level-0 stiffness matrix  $A_0 = [a_{0,i,j}] \in \mathbb{R}^{n_0 \times n_0}$  satisfies the Galerkin condition. Then  $A_0$  is clearly SPD and has the eigen-pairs*

$$\begin{aligned} \left[ \mathbf{v}_0^{(k)} \right]_i &= \sin(k\pi x_{0,i}), \quad 1 \leq i \leq n_0, \\ \lambda_0^{(k)} &= \frac{2}{h_0} (1 - \cos(k\pi h_0)), \end{aligned} \tag{21}$$

*for  $k = 1, \dots, n_0$ , where*

$$x_{0,i} = ih_0, \quad 0 \leq i \leq n_0 + 1.$$

*and*

$$h_0 = 2h_1 = \frac{1}{n_0 + 1}.$$