Problem 2 Overview B: There are 2 ways to solve $Wc = b$!!!!

1a) Solving $Wc = b$ is essentially a *change-of-basis* operation

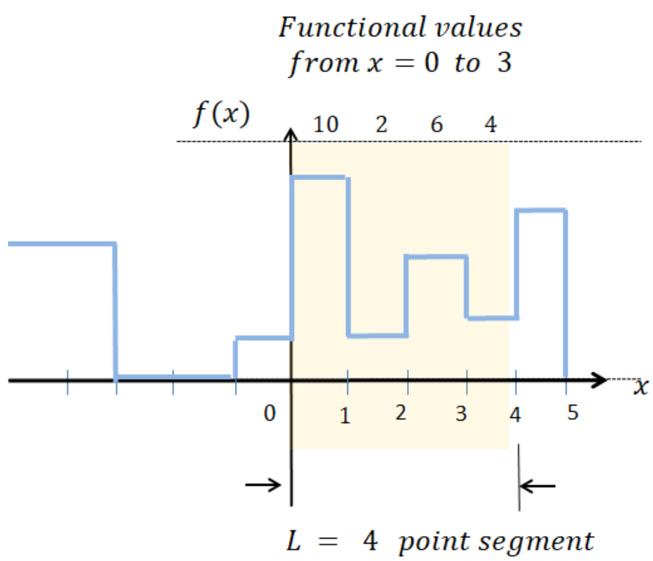


Figure 1: Our original wavelet problem, with a data word length of $L = 4$.

Recall our original 4-point Haar matrix problem: We would like to reconstruct a target vector b as a linear combo of the Haar wavelet basis:

$$b = c_0 \overrightarrow{w_0} + c_1 \overrightarrow{w_1} + c_2 \overrightarrow{w_2} + c_3 \overrightarrow{w_3}$$

In matrix form, this equation becomes our familiar friend:

$$\begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

$$b \qquad = \qquad W \qquad\qquad c$$

In the context of change-of-basis, we know from Lecture #8 that we have 2 "worlds" to contend with !

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

$$\underset{\left(\begin{smallmatrix}Home\ world\\ basis\end{smallmatrix}\right)}{I} \quad \underset{\begin{smallmatrix}Home\\ coords\end{smallmatrix}}{b} \quad = \quad \underset{\left(\begin{smallmatrix}Wavelet\ world\\ basis\end{smallmatrix}\right)}{W} \quad \underset{\begin{smallmatrix}Wavelet\\ coords\end{smallmatrix}}{c}$$

Hence, the wavelet coefficients c are just the corresponding coordinates of b in wavelet world. We can easily solve for :

$$c = W^{-1} b$$

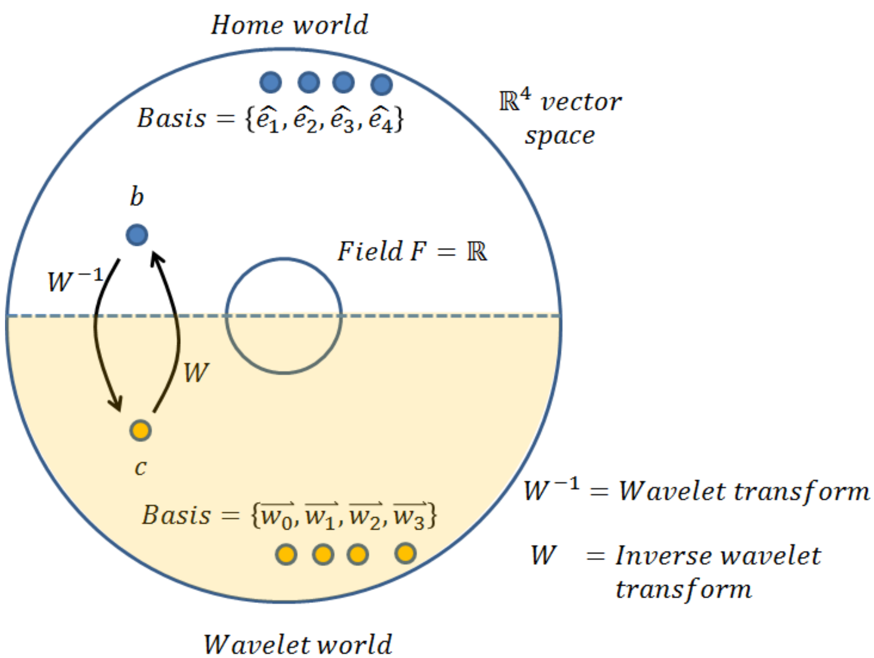
This is the "wavelet transform equation" !!!

Where:

$$W^{-1} = \text{ is called the } \begin{cases} Forward\ wavelet\ transform \\ The\ analysis\ filter\ \ in\ signals\ \&\ systems\ classes \end{cases}$$

You might wonder: Why did engineers and mathematicians assign W^{-1} as the "forward" wavelet transform and not W? The reason is that W^{-1} is the matrix that *shuttles our object b from the Home world → Wavelet world.*

In another words, W^{-1} is the engine that enables the change-of-basis operation from the raw data world → the compressed data world, *where the "compressed data" are the generalized Fourier coefficients c* (aka. "wavelet coefficients"). This can be seen in the crude abstract algebra diagram in Figure 2.



<u>Figure 2</u>: W^{-1} is called the <u>*forward avelet transform*</u> matrix because it acts on the home-world coordinate b (raw data) and shuttles it to the wavelet world coordinate c (compressed data).

1b) Solving $Wc = b$ is <u>*almost*</u> as easy as taking the transpose… but not quite!

To solve for c, we have to evaluate the wavelet transform equation:

$$c = W^{-1}\, b \qquad \rightarrow \qquad \begin{array}{l} Question\text{: Is it easy to evaluate } W^{-1}\ ? \\ Answer\text{:} \quad Yes\ !! \quad =) \end{array}$$

The column vectors are orthogonal… but not normal !

Let's take a look at why W^{-1} is easy to find. Suppose we suspect that W^{-1} was close to enjoying the same privilege as proper orthogonal matrices with orthonormal column vectors. We now ask the following question:

$$Is\ this\ true\text{:} \qquad W^T W = I\ ??$$

$$\boxed{Does \quad W^T W = I \ ??} \rightarrow \quad W^T W = \begin{bmatrix} \overrightarrow{w_0} \\ \hline \overrightarrow{w_1} \\ \hline \overrightarrow{w_2} \\ \hline \overrightarrow{w_3} \end{bmatrix} \left[\begin{array}{c|c|c|c} \overrightarrow{w_0} & \overrightarrow{w_1} & \overrightarrow{w_2} & \overrightarrow{w_3} \end{array} \right]$$

$$= \begin{bmatrix} \|\overrightarrow{w_0}\|^2 & & & \\ & \|\overrightarrow{w_1}\|^2 & & \\ & & \|\overrightarrow{w_2}\|^2 & \\ & & & \|\overrightarrow{w_3}\|^2 \end{bmatrix}$$

$$= \begin{bmatrix} 4 & & & \\ & 4 & & \\ & & 2 & \\ & & & 2 \end{bmatrix} \neq \begin{array}{l} Does\ not\ equal\ to\ the \\ identity\ matrix \quad = (\end{array}$$

But don't despair !! We can rescue this equation by multiplying a slick diagonal matrix on both sides :

$$\begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} W^T W = \underbrace{\begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} \begin{bmatrix} 4 & & & \\ & 4 & & \\ & & 2 & \\ & & & 2 \end{bmatrix}}$$

$$\begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} W^T W = I$$

Yes !!! =)

If we now apply a right-hand W^{-1} on both sides of the equation:

$$\begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} W^T \cancel{W} \cancel{W^{-1}} = I \, W^{-1}$$

Strang (3rd ed), Ch. 7.3,

page 386, middle of page

Then..... yay !! We have just found the expression for the inverse Haar matrix W^{-1}:

$$\boxed{W^{-1} = \begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} W^T} = \begin{bmatrix} \frac{1}{(norm)^2}\ values\ of\ each\ column \\ \\ located\ along\ the\ diagonal \end{bmatrix} W^T$$

- To find W^{-1} , a renormalization matrix is needed in front of W^T
- Our Haar matrix W is a **biorthogonal** matrix (contain orthogonal columns, but not unit vectors)

1c) Finding W^{-1} for larger basis sets:

You can use the same formula as before !

To find the inverse of any biorthogonal matrix W, you can actually use the same derivations from the previous page, where:

$$\begin{array}{l} \textit{If } W \textit{ is biorthogonal,} \\ \textit{then this must be true} \end{array} : \qquad W^T W = \begin{bmatrix} (norm)^2 \textit{ values of each column} \\ \\ \textit{located along the diagonal} \end{bmatrix}$$

$$\textit{This immediately implies:} \qquad W^{-1} = \begin{bmatrix} \frac{1}{(norm)^2} \textit{ values of each column} \\ \\ \textit{located along the diagonal} \end{bmatrix} W^T$$

Example: Find W^{-1} for the 8-point Haar matrix

If we had:

$$W = \begin{array}{c} \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & -1 & 0 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 0 & -1 & 0 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 1 & 0 & 0 & -1 & 0 \\ 1 & -1 & 0 & -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & -1 & 0 & 0 & 0 & -1 \end{bmatrix} \\ \begin{array}{cccccccc} \overrightarrow{w_0} & \overrightarrow{w_1} & \overrightarrow{w_2} & \overrightarrow{w_3} & \overrightarrow{w_4} & \overrightarrow{w_5} & \overrightarrow{w_6} & \overrightarrow{w_7} \end{array} \end{array}$$

The inverse of W is then:

$$W^{-1} = \begin{bmatrix} 1/8 & & & & & & & \\ & 1/8 & & & & & & \\ & & 1/4 & & & & & \\ & & & 1/4 & & & & \\ & & & & 1/2 & & & \\ & & & & & 1/2 & & \\ & & & & & & 1/2 & \\ & & & & & & & 1/2 \end{bmatrix} W^T$$

1d) The wavelet coefficients "c" $\approx$ Global averages and local differences within $\vec{b}$

You might be thinking: *OMG… can we solve the damn thing already !!* Yup – let's do it !

$$\begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & -1 & 0 \\ 1 & -1 & 0 & 1 \\ 1 & -1 & 0 & -1 \end{bmatrix} \cdot \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix}$$

$$b \quad = \quad W \quad c$$

Taking the inverse, we have:

$$c = W^{-1} b = \begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} W^T b$$

Biorthogonality !!

(requires a renormalization matrix in front of w^T…)

$$= \begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix}$$

Therefore, the wavelet coefficients (generalized Fourier coefficients in wavelet world !!) are:

$$c = \begin{bmatrix} 5.5 \\ 0.5 \\ 4 \\ 1 \end{bmatrix}$$

Yay ! This means our 4-point segment taken from $f(x)$ can be reconstructed using wavelets with ***different spatial frequencies*** :

$$b = \underbrace{5.5\ \overrightarrow{w_0}}_{} + \underbrace{0.5\ \overrightarrow{w_1}}_{} + \underbrace{4\ \overrightarrow{w_2} + 1\ \overrightarrow{w_3}}_{}$$

Super-coarse, global DC average within b

Coarse data variations across the entire b-vector

Fine data variations on the left / right sectors of the b-vector

A picture of this can be seen in Figure 6 !!

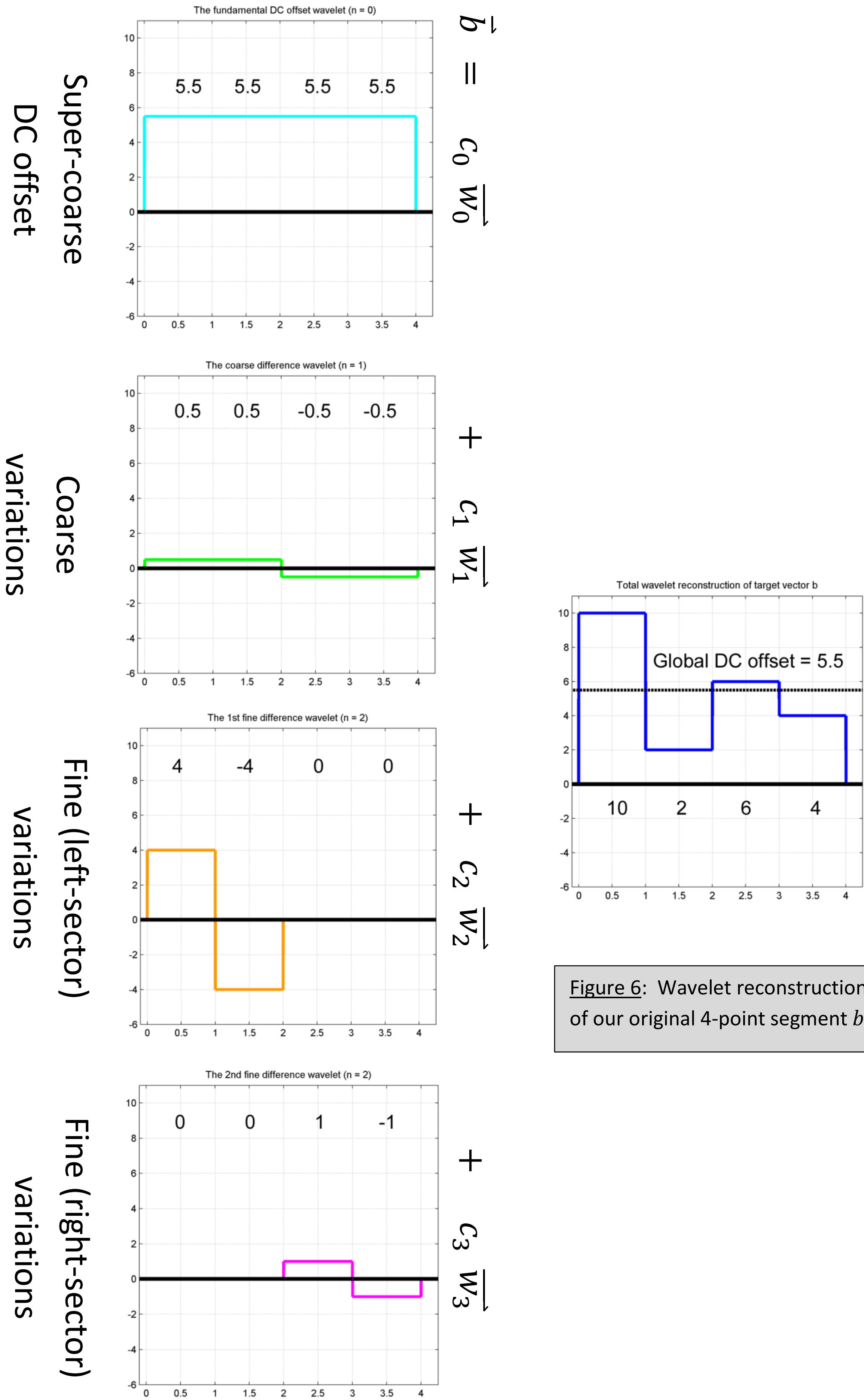


Figure 6: Wavelet reconstruction of our original 4-point segment b

2. There exists a "high-school" way of finding the Haar wavelet coefficients c

High school ? I've never heard of wavelets in high school ! Well, I'm gonna show you how easy it is !

2a) What happens at the fine-scale (J = 2):

Staring with b, first calculate the average of each adjacent pair and store them in a new vector:

$$\begin{bmatrix} f(0) \\ f(1) \\ f(2) \\ f(3) \end{bmatrix} = \begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix} \rightarrow \begin{bmatrix} 6 \\ 5 \\ \text{---} \\ \ \end{bmatrix}$$

$target\ b$

$$Top\ average\ bin = \frac{1}{2}\begin{pmatrix} Pairwise \\ sums \end{pmatrix}$$

$$= \frac{1}{2}\,[\,f(x) + f(x+1)\,]$$

Then, you take the ½ (differences) between each adjacent pair and place them in the bottom bin:

Fine (left) differences

Fine (right) differences

$$\begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix} \rightarrow \begin{bmatrix} 6 \\ 5 \\ \text{---} \\ 4 \\ 1 \end{bmatrix}$$

$target\ b$

$$Top\ average\ bin = \frac{1}{2}\begin{pmatrix} Pairwise \\ sums \end{pmatrix}$$

$$Bottom\ difference\ bin = \frac{1}{2}\begin{pmatrix} Pairwise \\ differences \end{pmatrix}$$

$$= \frac{1}{2}\,[\,f(x) - f(x+1)\,]$$

> Wait a minute…. Holy cow – We have just found the 2 fine-difference wavelet coefficients !!
>
> $$\begin{cases} c_2 = 4 & (\ J = 2\ level, \quad left-sector\ \ data\ variations) \\ c_3 = 1 & (\ J = 2\ level, \quad right-sector\ data\ variations) \end{cases}$$

By the way: If I had to express the top vs. bottom bin operations in matrix form, it would be:

Top bin ➡

Bottom bin ➡

$$\underbrace{\begin{bmatrix} 1/2 & 1/2 & & \\ & & 1/2 & 1/2 \\ \hline 1/2 & -1/2 & & \\ & & 1/2 & -1/2 \end{bmatrix}}_{\text{Top \& bottom bin transformation matrix}} \cdot \begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix} = \begin{bmatrix} 6 \\ 5 \\ \hline 4 \\ 1 \end{bmatrix}$$

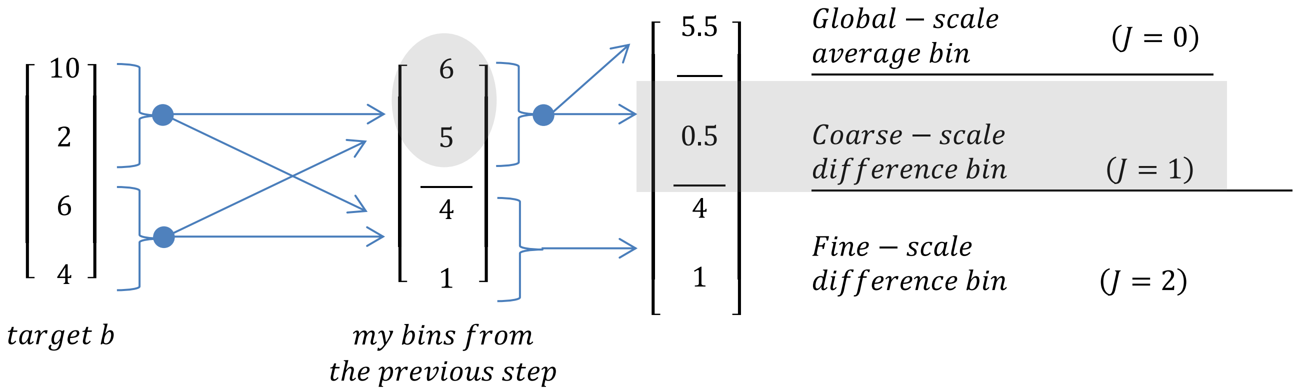
Top & bottom bin transformation matrix $\cdot$ b $=$ Contain partial answers for c !!

2b) What happens at the coarse-scale (J = 1) and global DC-scale (J = 0) :

Since we've already found the 2 fine-difference wavelet coefficients c_2 and c_3, let's leave the bottom bin alone and focus our attention on the top bin. We now apply the same operations again:

- Take the average of the top bin elements and place the answer in the top ½ of the top bin
- Take the ½ differences of the same elements and place the answer in the lower ½ of the top bin

> Cool ! – We have just found the coarse variation and the global DC wavelet coefficients !!
>
> $$\begin{cases} c_0 = 5.5 & (J = 0,\ \ global\ average\ value) \\ c_1 = 0.5 & (J = 1,\ \ coarse\ differences) \end{cases}$$



Now, the combinations of the current *and* the previous averaging / difference operations can also be written in a neat matrix form:

> <u>Identity matrix on lower-right corner means:</u>
>
> *Please don't messed with our hard-earned answers for* $c_2 = 4$ *and* $c_3 = 1$ *!!*

$$\underbrace{\begin{bmatrix} 1/2 & 1/2 & & \\ 1/2 & -1/2 & & \\ & & 1 & \\ & & & 1 \end{bmatrix}}_{J = 0 \text{ and } J = 1 \text{ (coarse) bin operations}} \cdot \underbrace{\begin{bmatrix} 1/2 & 1/2 & & \\ & & 1/2 & 1/2 \\ 1/2 & -1/2 & & \\ & & 1/2 & -1/2 \end{bmatrix}}_{J = 2 \text{ (fine) bin operations}} \cdot \underbrace{\begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix}}_{b} = \begin{bmatrix} 5.5 \\ 0.5 \\ 4 \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} c_0\ (J = 0) \\ c_1\ (J = 1) \\ c_2\ (J = 2) \\ c_3\ (J = 2) \end{bmatrix}}_{\text{Our complete set of wavelet coefficients !!}}$$

3c) Mystery solved: Haar basis = Really easy *multi-scale averages / differences*!

Last, but not least........ : What if we multiplied those 2 "high-school bin operation" matrices together ?

$$\underbrace{\begin{bmatrix} 1/2 & 1/2 & & \\ 1/2 & -1/2 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & 1/2 & & \\ & & 1/2 & 1/2 \\ 1/2 & -1/2 & & \\ & & 1/2 & -1/2 \end{bmatrix}}_{W^{-1}} = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 \\ 1/4 & 1/4 & -1/4 & -1/4 \\ 1/2 & -1.2 & 0 & 0 \\ 0 & 0 & 1/2 & -1/2 \end{bmatrix}$$

After staring at this matrix for a while, you'll realize the following is true:

$$= \begin{bmatrix} 1/4 & & & \\ & 1/4 & & \\ & & 1/2 & \\ & & & 1/2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

$$W^{-1} = \begin{pmatrix} biorthogonal \\ renormalization \\ matrix \end{pmatrix} \cdot W^T$$

Therefore, our high-school way of solving for c must be equivalent to :

$$\underbrace{\begin{bmatrix} 1/2 & 1/2 & & \\ 1/2 & -1/2 & & \\ & & 1 & \\ & & & 1 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & 1/2 & & \\ & & 1/2 & 1/2 \\ 1/2 & -1/2 & & \\ & & 1/2 & -1/2 \end{bmatrix}}_{W^{-1}} \cdot \underbrace{\begin{bmatrix} 10 \\ 2 \\ 6 \\ 4 \end{bmatrix}}_{b} = \underbrace{\begin{bmatrix} 5.5 \\ 0.5 \\ 4 \\ 1 \end{bmatrix}}_{c} = \begin{bmatrix} c_0 \ (J=0) \\ c_1 \ (J=1) \\ c_2 \ (J=2) \\ c_3 \ (J=2) \end{bmatrix}$$

$$\begin{matrix} We\ have\ now \\ proved\ that \end{matrix} : \boxed{\begin{Bmatrix} High-school \\ additions\ \& \\ subtractions \end{Bmatrix} = W^{-1}} = \begin{matrix} Taking\ the\ iverse\ of\ W \\ is\ uber\ easy\ !! \end{matrix} \quad =)$$

The main advantages of the high-school method

As you might guess, high-school binning method is preferred in modern implementations of the wavelet transform ver the formula

$$c \quad = \qquad W^{-1} \qquad b$$

$$= \left[\begin{array}{c} \frac{1}{(norm)^2} \; values\ of\ each\ column \\ \\ located\ along\ the\ diagonal \end{array} \right] W^T \quad b$$

This is because:

a) You don't have to waste memory trying building the matrix W

b) You don't have to waste read / write time by calculating W^T

c) You don't have to waste memory trying to calculate the blue renormalization matrix

d) You don't have to waste read / write time multiplying out the 3 above matrices

All you need for the high-school method are:

i) Write a loop

ii) Keep storing the "bottom bin differences" = The high-spatial frequency wavelet coefficients

iii) As soon as you have enough c_n's for your particular application, you can stop at any "J-level" you want.

Moral of the story:

- The simple, high-school level algorithm immediately elucidates the true meaning of the Haar wavelet coefficients c_n: They are just averages and differences from fine scale data variations $\rightarrow$ coarse scale data variations.

- When the target vector b becomes prohibitively large, people always use the high-school method to find the wavelet coefficients c_n because it is slightly faster if it's used in conjunction with sparse matrix algorithms !