Example Solutions for Classroom Assignment 3 (C3)

Problem 1: Discrete Fourier Transform

(a) This part of the exercise should incite some reflection about the meaning of high frequencies. A common misconception is that the value of the coefficients $\hat{f}_i$ determines which frequencies are high, but it doesn't. A coefficient representing a high frequency simply belongs to a basis function with high frequency and thus typically accounts for the amount of small detail changes in images.

The homework exercise H1, Problem 1 emphasises that the Fourier transform can be interpreted as a change of basis. The coefficients $f_0, \ldots, f_{M-1}$ of a 1-D signal of length M correspond to the canonical basis vectors $\boldsymbol{e}_0, \ldots, \boldsymbol{e}_{M-1}$. The discrete Fourier transform now expresses this signal in terms of M complex-valued basis vectors

$$\boldsymbol{b}_p \;:=\; \tfrac{1}{\sqrt{M}} \left(\exp\left(\tfrac{i2\pi p0}{M}\right), \exp\left(\tfrac{i2\pi p1}{M}\right), ..., \exp\left(\tfrac{i2\pi p(M-1)}{M}\right) \right)^\top$$
$$(p = 0, ..., M-1)$$

In Figure 1, the basis functions for the case $M = 8$ of this exercise are displayed. The highest frequency can be found in the middle of the signal and corresponds to $\hat{f}_4$. However, in practice, the origin, where the low frequencies are, is usually shifted to the image centre if we visualise the Fourier spectrum in terms of the coefficients magnitude. This applies in particular to the programming exercises dealing with the Fourier transform.

(b) If one sets the coefficients $\hat{f}_3$, $\hat{f}_4$ and $\hat{f}_5$ to zero and performs the back-transform, one obtains

$$\begin{aligned}&\left(13,\ 12+3\sqrt{2},\ 13+2\sqrt{2},\ 12,\ 7,\ 8-3\sqrt{2},\ 7-2\sqrt{2},\ 8\right)^\top\\ &\approx (13, 16.24, 15.83, 12, 7, 3.76, 4.27, 8)^\top .\end{aligned}$$

In comparison to the original signal, the suppression of the high frequencies has led to a considerably smoother result. However, at the

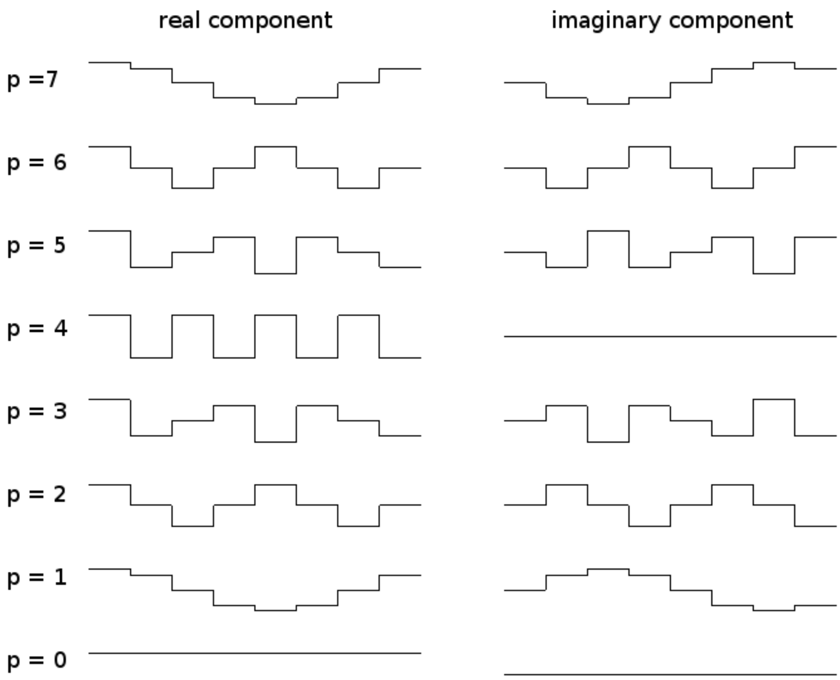

Figure 1: Real and imaginary components of the basis functions for the 1-D Fourier transform with signal length $M = 8$.

boundaries we notice rather obvious deviations from the original signal. This is due to the fact, that the Fourier Transform performs a periodic extension and thus the first term element of our signal is influenced by the last one and vice-versa.

Problem 2 (Image Pyramids)

(a) The *Gaussian pyramid* $\left\{\boldsymbol{v}^N, ..., \boldsymbol{v}^0\right\}$ of a signal $\boldsymbol{u} = (u_0, ..., u_{2^N-1})^\top$ is defined as

$$\begin{aligned} \boldsymbol{v}^N &:= \boldsymbol{u}, \\ \boldsymbol{v}^{k-1} &:= \boldsymbol{R}_k^{k-1}\boldsymbol{v}^k \quad (k = N, \ldots, 1), \end{aligned}$$

with

$$\boldsymbol{R}_k^{k-1} := \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & \ldots & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & \ldots & 0 & 0 & 0 & 0 \\ & & \ddots & \ddots & \ddots & \ddots & \ddots & & \\ 0 & 0 & 0 & 0 & \ldots & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & \ldots & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \qquad (k \geq 1)\ .$$

Applied to our signal $\boldsymbol{u} := (6,4,5,1,8,8,0,8)^\top$ we get

$$\begin{aligned} \boldsymbol{v}^3 &= \boldsymbol{u} = (6,4,5,1,8,8,0,8)^\top \\ \boldsymbol{v}^2 &= \boldsymbol{R}_3^2\boldsymbol{v}^3 \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 6 \\ 4 \\ 5 \\ 1 \\ 8 \\ 8 \\ 0 \\ 8 \end{pmatrix} \\ &= (5,3,8,4)^\top \\ \boldsymbol{v}^1 &= \boldsymbol{R}_2^1\boldsymbol{v}^2 \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 8 \\ 4 \end{pmatrix} \\ &= (4,6)^\top \\ \boldsymbol{v}^0 &= \boldsymbol{R}_1^0\boldsymbol{v}^1 \\ &= \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} \\ &= 5 \end{aligned}$$

(b) The *Laplacian pyramid* $\left\{\boldsymbol{w}^N, \ldots, \boldsymbol{w}^0\right\}$ of a signal $\boldsymbol{u} = (u_0, \ldots, u_{2^N-1})^\top$ with Gaussian pyramid $\left\{\boldsymbol{v}^N, \ldots, \boldsymbol{v}^0\right\}$ is defined as

$$\begin{aligned} \boldsymbol{w}^k &:= \boldsymbol{v}^k - \boldsymbol{P}_{k-1}^k \boldsymbol{v}^{k-1} \quad (k = N, \ldots, 1), \\ \boldsymbol{w}^0 &:= \boldsymbol{v}^0 \end{aligned}$$

with

$$\boldsymbol{P}_{k-1}^k := \begin{pmatrix} 1 & 0 & \ldots & 0 & 0 \\ 1 & 0 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ 0 & 1 & \ldots & 0 & 0 \\ & \ddots & \ddots & \ddots & \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 1 & 0 \\ 0 & 0 & \ldots & 0 & 1 \\ 0 & 0 & \ldots & 0 & 1 \end{pmatrix} \qquad (k \geq 1) .$$

Applied to our signal, we obtain

$$\boldsymbol{w}^0 = \boldsymbol{v}^0 = 5$$

$$\begin{aligned} \boldsymbol{w}^1 &= \boldsymbol{v}^1 - \boldsymbol{P}_0^1 \boldsymbol{v}^0 \\ &= \begin{pmatrix} 4 \\ 6 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} 5 = \begin{pmatrix} -1 \\ 1 \end{pmatrix} \end{aligned}$$

$$\begin{aligned} \boldsymbol{w}^2 &= \boldsymbol{v}^2 - \boldsymbol{P}_1^2 \boldsymbol{v}^1 \\ &= \begin{pmatrix} 5 \\ 3 \\ 8 \\ 4 \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} \\ &= (1, -1, 2, -2)^\top \end{aligned}$$

$$\begin{aligned} \boldsymbol{w}^3 &= \boldsymbol{v}^3 - \boldsymbol{P}_2^3 \boldsymbol{v}^2 \\ &= \begin{pmatrix} 6 \\ 4 \\ 5 \\ 1 \\ 8 \\ 8 \\ 0 \\ 8 \end{pmatrix} - \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 8 \\ 4 \end{pmatrix} \\ &= (1, -1, 2, -2, 0, 0, -4, 4)^\top \end{aligned}$$

(c) The Laplacian pyramid allows to *reconstruct* the original signal via the Gaussian pyramid:

$$\begin{aligned}
\boldsymbol{v}^0 &= \boldsymbol{w}^0 \\
\boldsymbol{v}^k &= \boldsymbol{w}^k + \boldsymbol{P}_{k-1}^k \boldsymbol{v}^{k-1} \quad (k = 1, \ldots, N), \\
\boldsymbol{u} &= \boldsymbol{v}^N
\end{aligned}$$

Let us now reconstruct the original signal from the Laplacian pyramid:

$$\begin{aligned}
\boldsymbol{v}^0 &= \boldsymbol{w}^0 = 5 \\
\boldsymbol{v}^1 &= \boldsymbol{w}^1 + \boldsymbol{P}_0^1 \boldsymbol{v}^0 \\
&= \begin{pmatrix} -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 1 \end{pmatrix} 5 = \begin{pmatrix} 4 \\ 6 \end{pmatrix} \\
\boldsymbol{v}^2 &= \boldsymbol{w}^2 + \boldsymbol{P}_1^2 \boldsymbol{v}^1 \\
&= \begin{pmatrix} 1 \\ -1 \\ 2 \\ -2 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} \\
&= (5, 3, 8, 4)^\top \\
\boldsymbol{u} = \boldsymbol{v}^3 &= \boldsymbol{w}^3 + \boldsymbol{P}_2^3 \boldsymbol{v}^2 \\
&= \begin{pmatrix} 1 \\ -1 \\ 2 \\ -2 \\ 0 \\ 0 \\ -4 \\ 4 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 5 \\ 3 \\ 8 \\ 4 \end{pmatrix} \\
&= (6, 4, 5, 1, 8, 8, 0, 4)^\top
\end{aligned}$$

(d) The resulting Laplacian pyramids contain redundant information. As one can see from the resulting vectors $\boldsymbol{w}^k$, the vectors contains doubled entries with $w_{2i+1}^k = -w_{2i}^k$. With this knowledge, the size of vector $\boldsymbol{w}^k$ can be reduced by factor $\frac{1}{2}$ by storing only all vector entries with even index. In our example, it suffices to store a vector $\boldsymbol{w} = (w_0^3, w_2^3, w_4^3, w_6^3, w_0^2, w_2^2, w_0^1, w_0^0)^\top = (1, 2, 0, -4, 1, 2, -1, 5)^\top$.