NOTE: Although this lecture is devoted to time filtering of high frequencies, which may be dominated by sampling errors, the same ideas can be applied just as well to space data. An example, is to "cleanly" smooth out high wavenumbers using the Lanczos filter in a high resolution topography data set in order to use it in a model.

Time filtering, low and high pass filters:

Each wavenumber k corresponds to a frequency $\nu_k = \frac{2\pi k}{T}$ and a period T / k. We can define a nondimensional frequency by multiplying the frequency by Δt : This nondimensional frequency $\nu_k \Delta t = \frac{2\pi k}{T}\Delta t = \frac{2\pi k}{N}$ can vary between 0 and π :

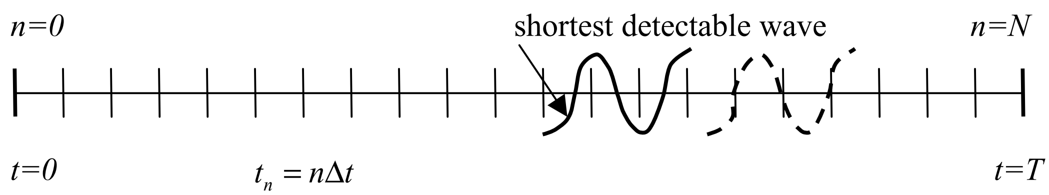
$$k = 0 \qquad \nu_k \Delta t = 0 \quad \text{zero frequency } (\infty \text{ period})$$

$$k_{\max} = \frac{N}{2} \quad \nu_{\max}\Delta t = \frac{2\pi}{N\Delta t}\frac{N}{2}\Delta t = \pi$$

The maximum frequency corresponds to the shortest period detectable,

$T_{\min} = 2\Delta t \qquad \nu_{\max} = \frac{2\pi}{2\Delta t} = \frac{\pi}{\Delta t}$, the Nyquist frequency.

Note also that at the Nyquist frequency, only the cosine can be detected, the sine is "invisible".



Clearly, this means that the wave with a period $2\Delta t$ is going to be poorly represented. Other short waves (with periods between $2\Delta t$ and $4\Delta t$) will also be distorted by the sampling every Δt . We will see later that there are other problems with the high frequency waves, including "aliasing". This leads to the desirability of applying low pass filters to the time series that smooth out the $2\Delta t$ waves as well as other short waves. For example, if we have twice daily data, the diurnal cycle is very poorly represented, we need to have data at least 4 times a day or more to represent the diurnal cycle. As a result, it is common to filter (smooth) time series so that the unresolved (high frequency) waves are not present.

In order to understand the impact (response) of different time filters, consider for simplicity a time series with a single (pure) frequency:

$$f_n = C_k \exp\left(ik \frac{2\pi}{N\Delta t} n\Delta t \right) = C_k \exp\left(i\nu_k \Delta t\, n \right) \qquad (1)$$

We now want to apply a time filter to reduce the amplitude of the $2\Delta t$ (highest frequency) component. There are a number of possible filters:

1) A good filter for a time series is the following:

$$\overline{f}_n = \frac{1}{2}\left(\frac{f_{n-1} + f_n}{2} + \frac{f_n + f_{n+1}}{2} \right) = \frac{f_{n-1} + 2f_n + f_{n+1}}{4}$$

This means that we replace the time series with a new, filtered one with the average indicated above.

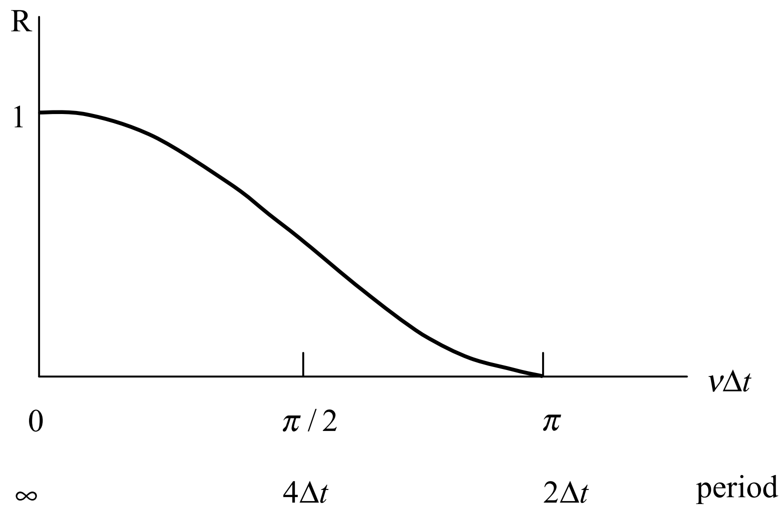
In order to understand the response, we apply the filter to the single frequency series (1) above:

$$\overline{f}_n = C_k e^{\left(i\nu_k \Delta t\, n\right)} . \frac{e^{\left(-i\nu_k \Delta t\right)} + 2 + e^{\left(i\nu_k \Delta t\right)}}{4} = f_k^n \frac{1 + \cos\left(\nu_k \Delta t\right)}{2}$$

So, the response of the filter for each frequency is

$$R(\nu_k) = \frac{1+\cos(\nu_k \Delta t)}{2}$$

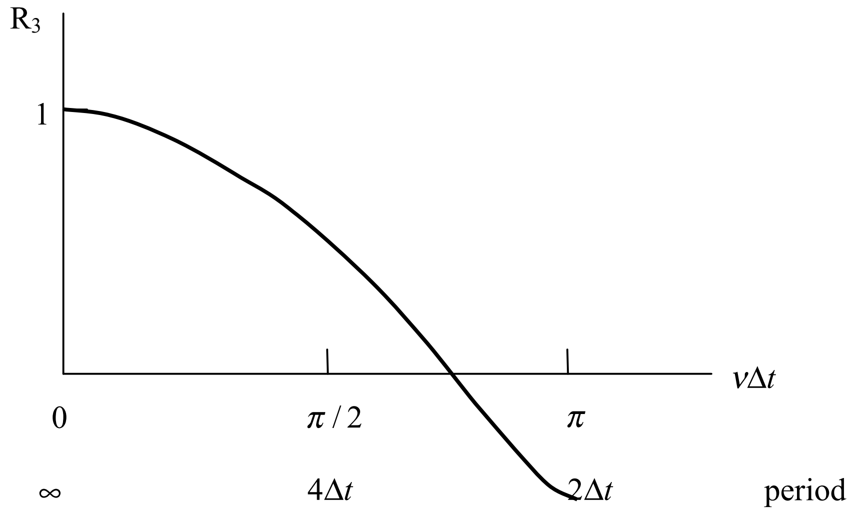
This filter has the nice property that it smoothes out the high frequencies, eliminating completely the $2\Delta t$ frequency:



2) Another commonly used time filter is the running mean:

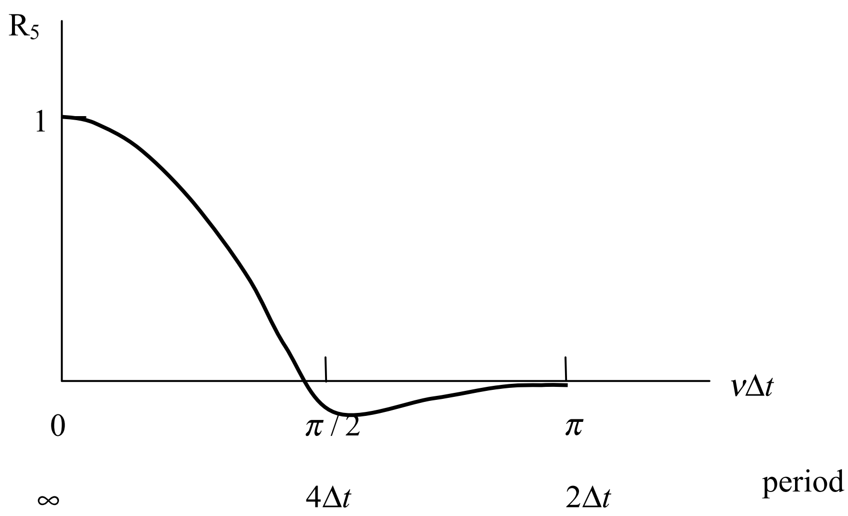
$$\overline{f}_{n,k}^{\;3} = \frac{f_{n-1}+f_n+f_{n+1}}{3} = C_k e^{(i\nu_k \Delta t\, n)} . \frac{e^{(-i\nu_k \Delta t)}+1+e^{(i\nu_k \Delta t)}}{3} = f_{n,k}\frac{1+2\cos(\nu_k \Delta t)}{3}$$

Not so good: for $T = 2\Delta t$, waves are multiplied by -1/3!

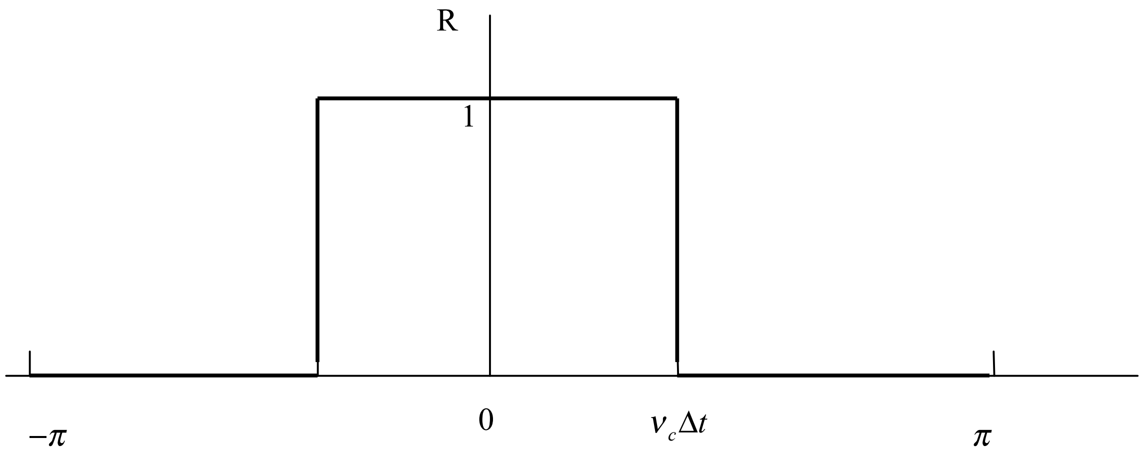


3) A 5-point running mean: better but not perfect:

$$\overline{f}_{n,k}^{\,5} = f_{n,k}\,\frac{1+2\cos\left(\nu_k\Delta t\right)+2\cos\left(2\nu_k\Delta t\right)}{5}$$



A perfect "low pass filter" would have a response equal to 1 for low frequencies lower than a critical frequency ν_c , and zero for higher frequencies:



In general, a filter has the form $\bar{f}^n = \sum_{l=-\infty}^{\infty} w_l f_{n-l}$

The response $R(\nu) = \dfrac{\bar{f}_\nu}{f_\nu}$ of such a general filter for a time series $f_\nu^n = e^{i\nu\Delta t n}$ with a single frequency ν is then

$$R(\nu) = \sum_{l=-\infty}^{\infty} w_l e^{i\nu\Delta t l}$$

If we choose the response (shape) of a filter $R(\nu)$, then we can obtain the weights through a Fourier transform of $R(\nu)$:

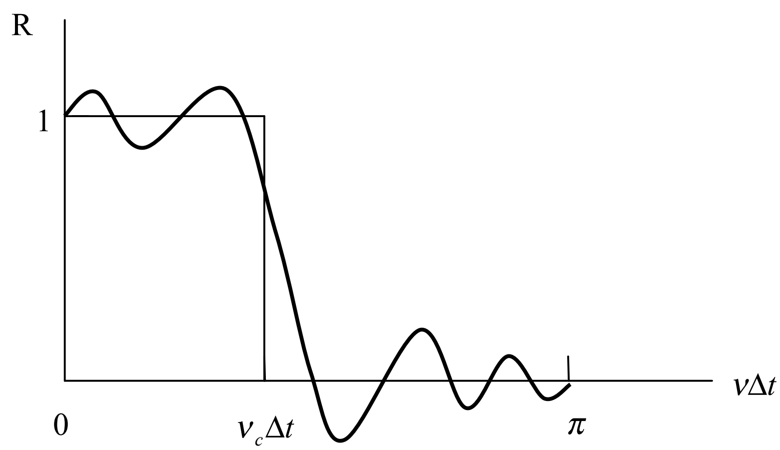
$$w_l = \frac{1}{2\pi}\int_{-\pi}^{\pi} R(\nu)e^{-i\nu\Delta t l}d(\nu\Delta t)$$

For the ideal low pass filter with a critical frequency $\nu_c\Delta t$,

$$w_l = \frac{1}{2\pi}\int_{-\nu_c\Delta t}^{\nu_c\Delta t} e^{-i(\nu\Delta t)l}d(\nu\Delta t) = \frac{1}{2\pi}\left(\frac{1}{-il}\right)\left[e^{-i\nu_c\Delta t l} - e^{i\nu_c\Delta t l}\right]$$

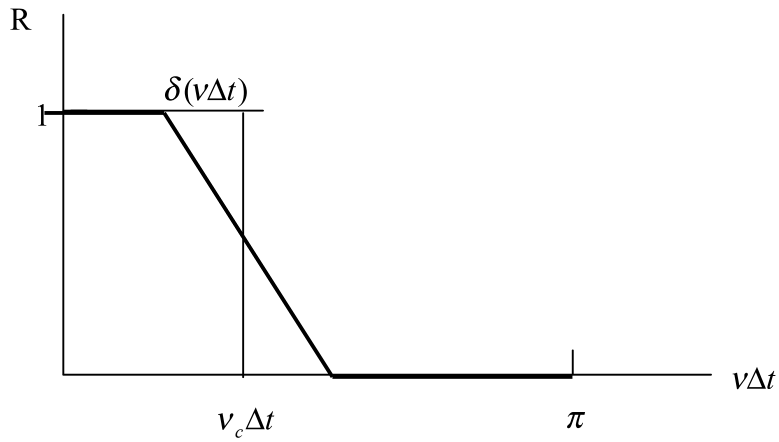
or $w_l = \dfrac{\sin(\nu_c\Delta t l)}{\pi l}$

These are the weights for an ideal low pass filter (with $l = -\infty,\dots,0,\dots,\infty$). However, in practice, such a filter can only have a finite number of terms L. The result of using a finite number of terms for a discontinuous function is that the Fourier representation will overshoot and undershoot (Gibbs phenomenon):



For this reason, Lanczos introduced a different “ideal” filter (**the Lanczos filter**), that instead of a step function, has a ramp around the critical frequency, with a width that depends on the number of terms L of the filter:

$$\delta(\nu\Delta t) = \frac{\pi}{L}$$



The Lanczos low pass filter, with essentially zero overshooting, is given by

$$\overline{w}_l^{LP} = w_l \bullet \frac{\sin(\pi l / L)}{\pi l / L} = \frac{\sin(\nu_c \Delta t l)}{\pi l} \frac{\sin(\pi l / L)}{\pi l / L}.$$

(Note that for $l = 0$, $\overline{w}_0^{LP} = \dfrac{\nu_c \Delta t}{\pi}$ using L’Hopital rule)

This filter has an almost perfect low-pass response. (Reference: Duchon, 1979, J. Applied Meteorology). Another filter used in digital initialization is the Dolph-Tchebychev filter (Lynch, 1997, MWR), which has slightly different properties than the Lanczos filter (see also Kalnay, 2003, p 194 on).

In summary, to apply a Lanczos Low-Pass filter (smoother) to a time series sampled at constant intervals Δt :

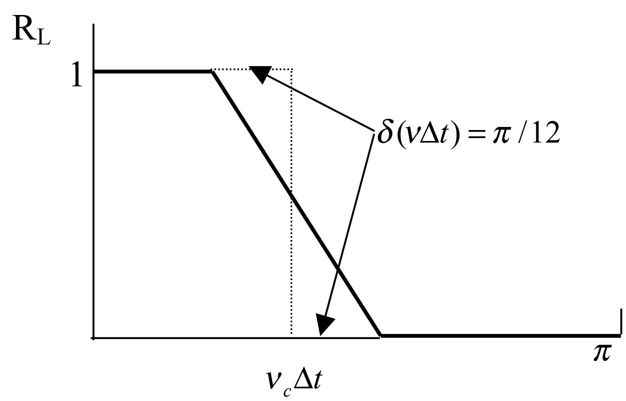
1) Choose the critical period, for example, daily data with T_c= 10 days.

Then, $\nu_c \Delta t = \frac{2\pi}{10\text{days}} \cdot 1\text{day} = \frac{\pi}{5}$

2) Choose the number of terms in the filter $2L+1$ (i.e., $l = -L, \ldots, 0, \ldots, L$)

This determines the width of the filter $\delta(\nu \Delta t) = \frac{\pi}{L}$

For example, if L=12 (25 term filter):

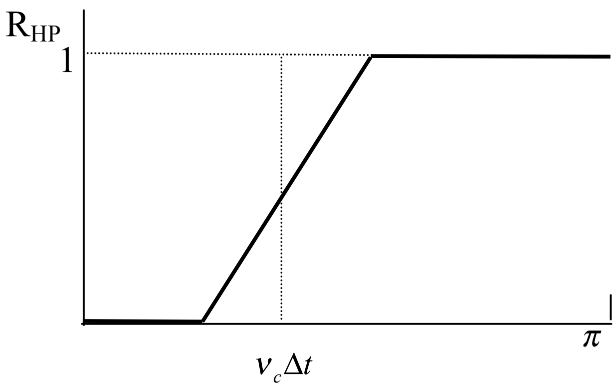


The low-pass time series becomes

$$\overline{f}_n^{LP} = \sum_{l=-12}^{12} \overline{w}_l^{LP} f_{n-l} \quad \text{with}$$

$$\overline{w}_l^{LP} = \frac{\sin\left(\frac{\pi}{5}l\right)}{\pi l} \cdot \frac{\sin\left(\frac{\pi}{12}l\right)}{\frac{\pi}{12}l}.$$

A **high pass filter** (filters out low frequencies and retains high frequencies) has a response $1 - \overline{R}^{LP}$

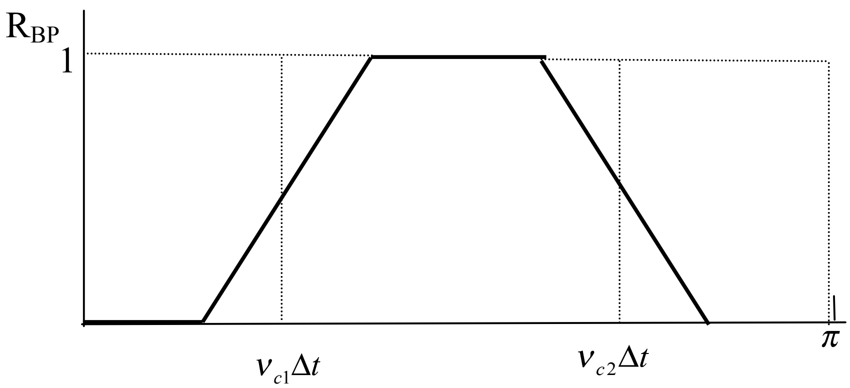


$$\overline{f}_n^{HP} = f_n - \overline{f}_n^{LP} \quad \text{or} \quad \overline{f}_n^{HP} = \sum_{l=-L}^{L} \overline{w}_l^{HP} f_{n-l}$$

with

$$\overline{w}_0^{HP} = 1 - \overline{w}_0^{LP}, \quad \overline{w}_l^{HP} = -\overline{w}_l^{LP}$$

A **band-pass filter** is constructed choosing two critical frequencies (or periods):



$$\bar{f}_n^{BP} = \sum_{l=-L}^{L} \bar{w}_l^{BP} f_{n-l}$$

with

$$\bar{w}_l^{BP} = \left[\frac{\sin(\nu_{c2}\Delta t l)}{\pi l} - \frac{\sin(\nu_{c1}\Delta t l)}{\pi l} \right] \frac{\sin(\pi l / L)}{\pi l / L}$$

Again, these ideas can be used to do space filtering of noisy data sets just as well as time filtering.