

Course material

Course:

Digital Signal Processing II

Video:

2.2.4.a Filter specifications

Concepts (extracted from automatically generated subtitles):

Filter design problem. Pass band. Frequency response. Form of a rational transfer function. Stop band. Important case. Final design. Small error tolerances. Filter of high order. Coefficients of the polynomials. Realizable filters. Gradual decay of the frequency response. Pass bands. Stop bands. Cut-off frequency.



[to video sequence search](#)
(within Digital Signal Processing II.)



[to video](#)

Center for Digital Education. More educational support material here:

<https://exoset.epfl.ch/booc/>

The filter design problem

You are given a set of requirements:

- ▶ frequency response: passband(s) and stopband(s)
- ▶ phase: overall delay, linearity
- ▶ some limit on computational resources and/or numerical precision

You must determine N , M , a_k 's and b_k 's in

$$H(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_{M-1} z^{-(M-1)}}{a_0 + a_1 z^{-1} + \dots + a_{N-1} z^{-(N-1)}}$$

in order to best fulfill the requirements

So let's review the filter design problem and suppose you are given a set of requirements, such as a frequency response specified in terms of where you want the pass band to be and where you want the stop band to be. At the same time you could have multiple stop bands and multiple pass bands. You will probably have some phase requirements as well, namely in the form of whether linear phase is necessary or not. And most often you will have some limit on the computational resources you can use. We are interested in realizable filters, therefore our final design will be expressed in the form of a rational transfer function.

notes

summary

0m 0s



The filter design problem

You are given a set of requirements:

- ▶ frequency response: passband(s) and stopband(s)
- ▶ phase: overall delay, linearity
- ▶ some limit on computational resources and/or numerical precision

You must determine N , M , a_k 's and b_k 's in

$$H(z) = \frac{b_0 + b_1 z^{-1} + \dots + b_{M-1} z^{-(M-1)}}{a_0 + a_1 z^{-1} + \dots + a_{N-1} z^{-(N-1)}}$$

in order to best fulfill the requirements



And our problem becomes finding the degree of the numerator and of the denominator,

notes

summary

0m 29s



Practical limitations

- ▶ passband/stopband transitions cannot be infinitely sharp
⇒ use *transition bands*
- ▶ magnitude response cannot be constant over an interval
⇒ specify *magnitude tolerances over bands*
- ▶ in general:
 - smaller transition bands ⇒ higher filter order
 - smaller error tolerances ⇒ higher filter order
 - higher filter order ⇒ more expensive, larger delay

as well as the coefficients of the polynomials involved, in order to best fulfill the requirements. So for example, if we want to design a low pass, the specifications could be like these. We have a cut-off frequency, we have a desired value for the pass band and a desired value for the stop band. Now these specifications seem to indicate an ideal characteristic that we know we cannot attain in practice. And therefore we will try to fulfill these requirements by taking explicitly into account what we know about what is achievable with a realizable filter. In particular, we know that the pass band and stop band transitions cannot be infinitely sharp. And therefore we will use some transition bands to allow for a gradual decay of the frequency response from pass band to stop band.

notes

summary

0m 37s



Practical limitations

- ▶ passband/stopband transitions cannot be infinitely sharp
⇒ use *transition bands*
- ▶ magnitude response cannot be constant over an interval
⇒ specify *magnitude tolerances over bands*
- ▶ in general:
 - smaller transition bands ⇒ higher filter order
 - smaller error tolerances ⇒ higher filter order
 - higher filter order ⇒ more expensive, larger delay

Also, the magnitude response cannot be constant over an interval, we will see shortly why. And therefore we will have to specify some tolerances over the pass bands and stop bands,

notes

summary

1m 25s



Practical limitations

- ▶ passband/stopband transitions cannot be infinitely sharp
⇒ use *transition bands*
- ▶ magnitude response cannot be constant over an interval
⇒ specify *magnitude tolerances over bands*
- ▶ in general:
 - smaller transition bands ⇒ higher filter order
 - smaller error tolerances ⇒ higher filter order
 - higher filter order ⇒ more expensive, larger delay

within which we allow the frequency response to move. In general, the lesson is the following.

notes

summary

1m 37s



Practical limitations

- ▶ passband/stopband transitions cannot be infinitely sharp
⇒ use *transition bands*
- ▶ magnitude response cannot be constant over an interval
⇒ specify *magnitude tolerances over bands*
- ▶ in general:
 - smaller transition bands ⇒ higher filter order
 - smaller error tolerances ⇒ higher filter order
 - higher filter order ⇒ more expensive, larger delay

If we want very small transition bands, we will need to use a filter of high order. And similarly, if we want small error tolerances, we need a high order filter. Now high order means a high polynomial degree either at the numerator or at the denominator of the transfer function, which means that we will need to use more computational power to implement that filter and that the delay introduced by the filter in a causal realization will be larger.

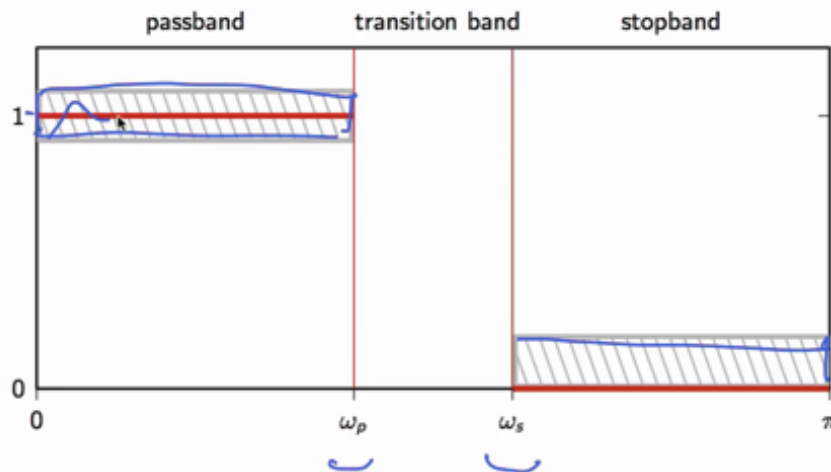
notes

summary

1m 47s



Example: realistic lowpass specs



We can plot the realistic low pass specifications graphically as follows. Instead of having one cutoff frequency, we have a transition band between a frequency ω_p , which specifies the end of the pass band, to a frequency ω_s , which specifies the beginning of the stop band. Instead of having one desired value for stop band and pass band, we have tolerance regions within which the frequency response can wiggle.

notes

summary

2m 13s



Why we can't have a flat response

$$H(z) = B(z)/A(z), \quad \text{with } A \text{ and } B \text{ polynomials}$$

$$\begin{aligned} H(e^{j\omega}) = c \text{ over an interval} &\Rightarrow B(z) - cA(z) = 0 \text{ over an interval} \\ &\Rightarrow B(z) - cA(z) \text{ has an infinite number of roots} \\ &\Rightarrow B(z) - cA(z) = 0 \text{ for all values of } z \\ &\Rightarrow H(e^{j\omega}) = c \text{ over the entire } [-\pi, \pi] \text{ interval.} \end{aligned}$$

The fact that we cannot have an arbitrarily sharp transition from pass band to stop band should be sufficiently clear. The transfer function is a rational transfer function, and as such it cannot have a discontinuity point. The reason why we cannot have a flat response, i.e. a response that is identical to a constant over a certain interval,

notes

summary

2m 46s



Why we can't have a flat response

$$H(z) = B(z)/A(z), \quad \text{with } A \text{ and } B \text{ polynomials}$$

$H(e^{j\omega}) = c$ over an interval $\Rightarrow B(z) - cA(z) = 0$ over an interval
 $\Rightarrow B(z) - cA(z)$ has an infinite number of roots
 $\Rightarrow B(z) - cA(z) = 0$ for all values of z
 $\Rightarrow H(e^{j\omega}) = c$ over the entire $[-\pi, \pi]$ interval.

maybe requires a little bit more of an explanation. So again, we start from a rational transfer function and now suppose that the frequency response,

notes

summary

3m 1s



Why we can't have a flat response

$$H(z) = B(z)/A(z), \quad \text{with } A \text{ and } B \text{ polynomials}$$

$$\begin{aligned} \underbrace{H(e^{j\omega}) = c \text{ over an interval}} &\Rightarrow B(z) - cA(z) = 0 \text{ over an interval} \\ &\Rightarrow B(z) - cA(z) \text{ has an infinite number of roots} \\ &\Rightarrow B(z) - cA(z) = 0 \text{ for all values of } z \\ &\Rightarrow H(e^{j\omega}) = c \text{ over the entire } [-\pi, \pi] \text{ interval.} \end{aligned}$$

$H(e^{j\omega}) = c$ is constant over an interval, no matter how small that interval. Well, in that case, the Z-transform will be constant over an interval as well, and therefore we can write that the denominator minus the constant times the numerator is identically zero over an interval.

notes

summary

3m 15s



Why we can't have a flat response

$$H(z) = B(z)/A(z), \quad \text{with } A \text{ and } B \text{ polynomials}$$

$$\begin{aligned} \underbrace{H(e^{j\omega}) = c \text{ over an interval}} &\Rightarrow \underbrace{B(z) - cA(z) = 0 \text{ over an interval}} \\ &\Rightarrow B(z) - cA(z) \text{ has an infinite number of roots} \\ &\Rightarrow B(z) - cA(z) = 0 \text{ for all values of } z \\ &\Rightarrow H(e^{j\omega}) = c \text{ over the entire } [-\pi, \pi] \text{ interval.} \end{aligned}$$

But this guy here, $B(z) - cA(z)$, is a polynomial in z , so if it is constant over an interval, it has an infinite number of roots over that interval.

notes

summary

3m 28s



Why we can't have a flat response

$$H(z) = B(z)/A(z), \quad \text{with } A \text{ and } B \text{ polynomials}$$

$$\begin{aligned} H(e^{j\omega}) = c \text{ over an interval} &\Rightarrow B(z) - cA(z) = 0 \text{ over an interval} \\ &\Rightarrow B(z) - cA(z) \text{ has an infinite number of roots} \\ &\Rightarrow B(z) - cA(z) = 0 \text{ for all values of } z \\ &\Rightarrow H(e^{j\omega}) = c \text{ over the entire } [-\pi, \pi] \text{ interval.} \end{aligned}$$

But we know from the fundamental theorem of algebra that if a polynomial has an infinite number of roots,

notes

summary

3m 45s



Why we can't have a flat response

$$H(z) = B(z)/A(z), \quad \text{with } A \text{ and } B \text{ polynomials}$$

$$\begin{aligned} H(e^{j\omega}) = c \text{ over an interval} &\Rightarrow B(z) - cA(z) = 0 \text{ over an interval} \\ &\Rightarrow B(z) - cA(z) \text{ has an infinite number of roots} \\ &\Rightarrow B(z) - cA(z) = 0 \text{ for all values of } z \\ &\Rightarrow H(e^{j\omega}) = c \text{ over the entire } [-\pi, \pi] \text{ interval.} \end{aligned}$$

it is identically zero over the entire complex plane, which means that if the Fourier transform is constant over an interval, no matter how small that interval, it will be constant over the entire frequency range. As a consequence, you can think of the frequency response of a filter as a shark. It must always move and it can never be at rest. An important case is what we call the equiripple error,

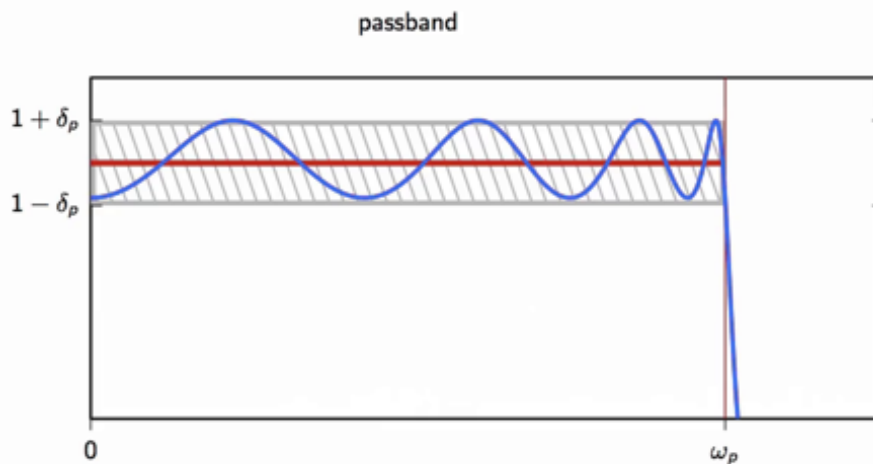
notes

summary

3m 49s



Important case: equiripple error



where the error, for instance, in this case in the pass band, oscillates between a maximum and a minimum, and the local extrema of the frequency response coincides with the upper and lower limit of the tolerance region. Once the specs are in place, the three big questions are,

notes

summary

4m 14s



The big questions

- ▶ IIR or FIR?
- ▶ how to determine the coefficients?
- ▶ how to evaluate the performance?

well first, are we going to design an IIR or an FIR?

notes

summary

4m 31s



IIRs: pros and cons

Pros:

- ▶ computationally efficient
- ▶ strong attenuation easy
- ▶ good for audio

Cons:

- ▶ stability issues
- ▶ difficult to design for arbitrary response
- ▶ nonlinear phase

And once we have answered that question, how will we determine the coefficients of the transfer function and how do we evaluate the performance of the filter? In order to answer the first question, we have to consider the pros and the cons of one design versus another. IIRs are computationally efficient. They can achieve a very strong attenuation in the stop band rather easily, and they're good for audio because they can achieve a monotonic characteristic in the pass band. On the other hand, they might have stability issues, especially in numerical implementations that are prone to overflow or underflow. They're difficult to design for arbitrary responses, namely a low pass or high pass characteristic is easy to obtain, but it's not so easy to obtain an arbitrary response with an IIR,

notes

summary

4m 37s



IIRs: pros and cons

Pros:

- ▶ computationally efficient
- ▶ strong attenuation easy
- ▶ good for audio

Cons:

- ▶ stability issues
- ▶ difficult to design for arbitrary response
- ▶ nonlinear phase

and they have nonlinear phase.

notes

summary

5m 22s



FIRs: pros and cons

Pros:

- ▶ always stable
- ▶ optimal design techniques exist
- ▶ can be designed with linear phase

Cons:

- ▶ computationally much more expensive
- ▶ may "sound" harsh

FIRs, on the other hand, they're always stable, they can be designed optimally, as we will see later, and they can be designed with linear phase, which is a great advantage, especially for communication systems. On the other hand, they're computationally more expensive, and acoustically they might sound harsh sometimes. In general, finding the degree and the coefficients of the polynomials involved in the transfer function

notes

summary

5m 25s



The design methods

- ▶ finding N , M , a_k 's and b_k 's from specs is a hard nonlinear problem
- ▶ established methods:
 - IIR: conversion of analog design
 - FIR: optimal minimax filter design

is a very hard nonlinear problem. In the case of IIR filters, there is a vast body of literature on filter design

notes

summary

5m 45s



The design methods

- ▶ finding N , M , a_k 's and b_k 's from specs is a hard nonlinear problem
- ▶ established methods:
 - IIR: conversion of analog design
 - FIR: optimal minimax filter design

that dates back from the days of discrete electronics that has been converted to the problem of designing digital filters. FIR filters, on the other hand, are purely digital entities that do not exist in the analog world,

notes

summary

5m 53s



The design methods

- ▶ finding N , M , a_k 's and b_k 's from specs is a hard nonlinear problem
- ▶ established methods:
 - IIR: conversion of analog design
 - FIR: optimal minimax filter design

and so design techniques for FIR filters have been developed from scratch since the beginning of digital signal processing, and they have reached their apex in the early 70s with Parks McClellan's Optimo filter design technique.

notes

summary

6m 6s

