Filter Design

Daniel P. Palomar
Hong Kong University of Science and Technology (HKUST)

ELEC5470 - Convex Optimization
Fall 2018-19, HKUST, Hong Kong

Outline of Lecture

- FIR filters
- Chebychev design
- Lowpass filter design
- Filter magnitude specification design
- Log-Chebychev magnitude specification design
- Equalizer design
- Summary

(Acknowledgement to Stephen Boyd for material for this lecture.)

FIR Filters

- The input-output relationship for a finite-impulse response (FIR) filter is

$$y(t) = \sum_{\tau=0}^{n-1} h_\tau x(t-\tau)$$

  where

  - $x(t)$ is the real-valued input sequence
  - $y(t)$ is the real-valued output sequence
  - h_i are the real-valued filter coefficients
  - n is the filter order or length.

- Observe that the output is a linear function of the input.

- The *FIR filter frequency response* is

$$\begin{aligned} H(\omega) &= h_0 + h_1 e^{-j\omega} + \cdots + h_{n-1} e^{-j(n-1)\omega} \\ &= \sum_{t=0}^{n-1} h_t \cos t\omega - j \sum_{t=0}^{n-1} h_t \sin t\omega. \end{aligned}$$

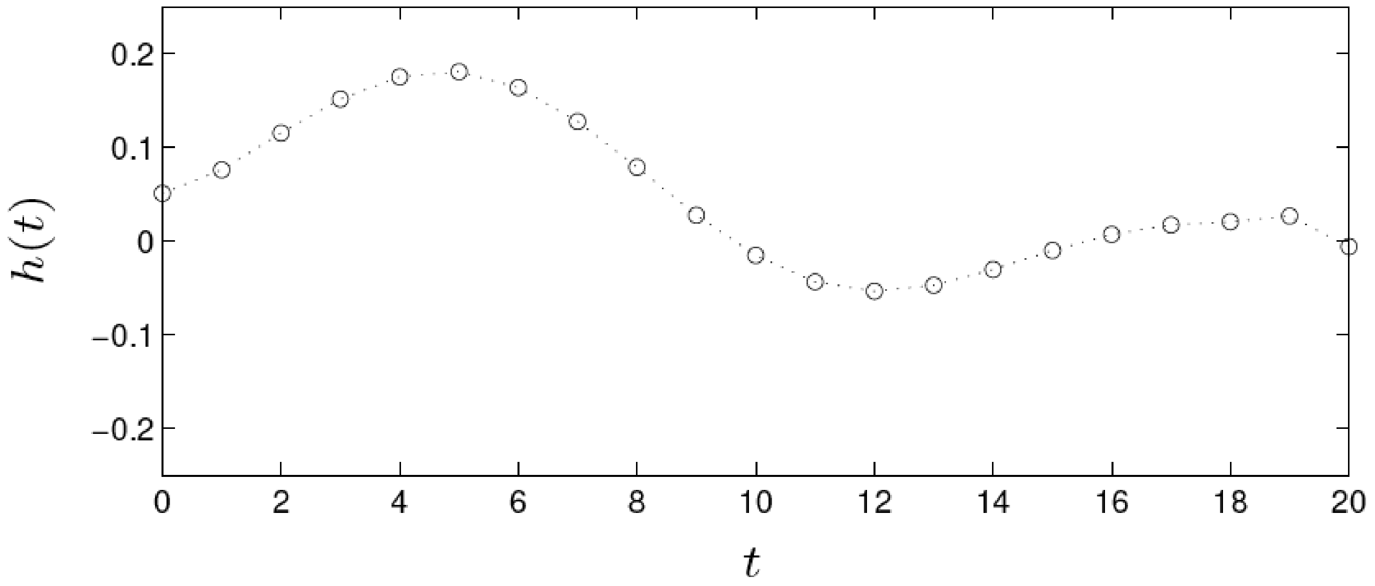
- Observations:
  - $H(\omega)$ is complex-valued
  - $H(\omega)$ is periodic and conjugate symmetric, so only need to specify for $0 \leq \omega \leq \pi$.
  - $H(\omega)$ is a linear function of the filter coefficients.

- The *FIR filter design problem* is to design $\mathbf{h}$ such that it and $H(\omega)$ satisfy/optimize some specifications.

FIR Example

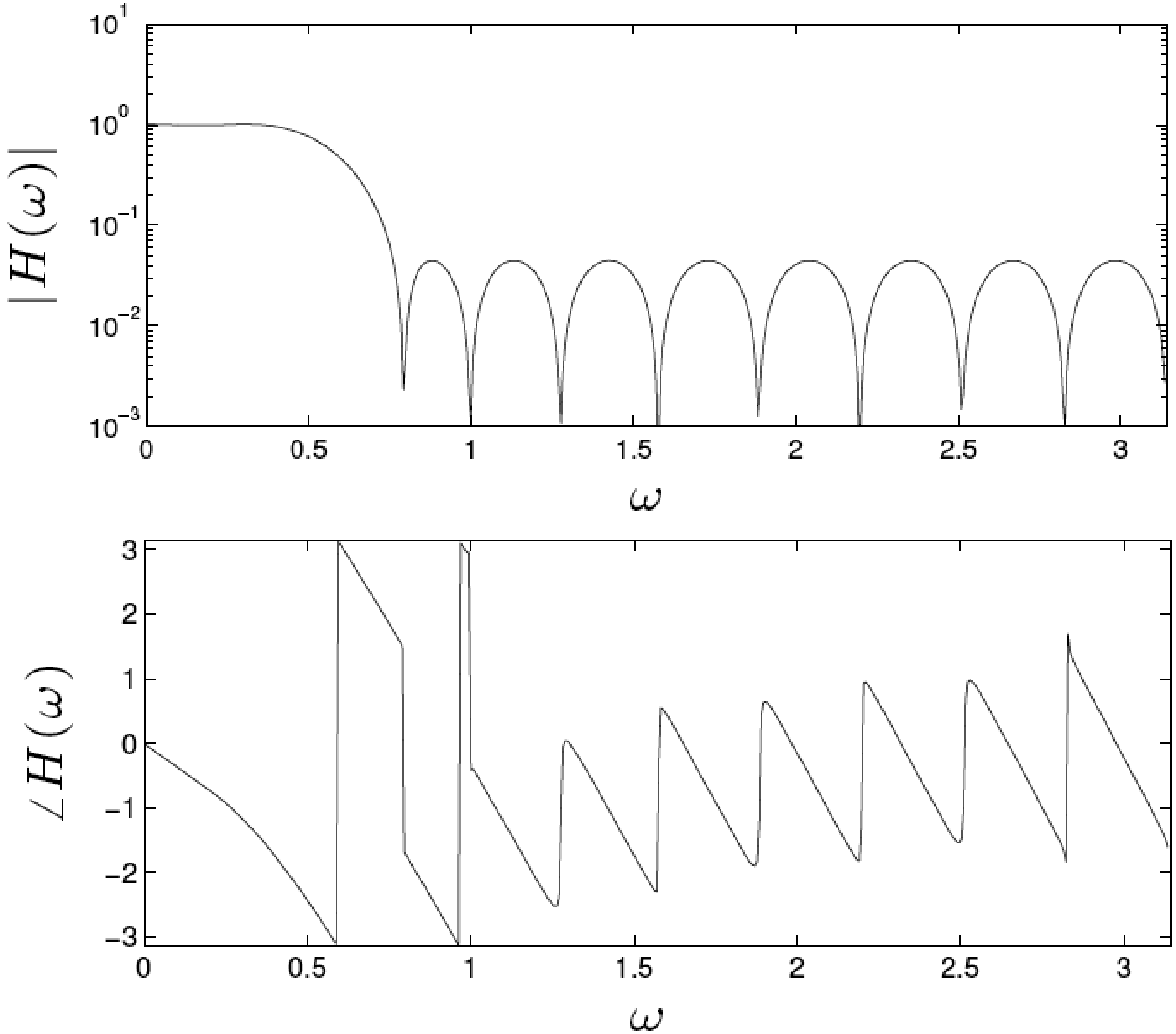
- Example of a lowpass FIR filter of order $n = 21$.

- The impulse response $\mathbf{h}$ is



from which it is difficult to infer the filtering capabilities and properties.

- The frequency response magnitude and phase, $|H(\omega)|$ and $\angle H(\omega)$, are



Chebychev Design

- The problem formulation is

$$\underset{\mathbf{h}}{\text{minimize}} \max_{\omega \in [0,\pi]} \left| H(\omega) - H_{\text{des}}(\omega) \right|$$

  where

  – $\mathbf{h}$ is the optimization variable (recall that $H(\omega)$ is linear in $\mathbf{h}$)
  – $H_{\text{des}}(\omega)$ is the desired transfer function.

- This Chebychev formulation is a semi-infinite convex problem. (Why?)

- We can add constraints while keeping the convexity such as $|h_i| \leq 1$.

- In practice, to deal with the infinite set of frequencies, we discretize:

$$\underset{\mathbf{h}}{\text{minimize}} \ \max_{k=1,\cdots,m} \left| H(\omega_k) - H_{\text{des}}(\omega_k) \right|$$

  where

  – sample points $0 \le \omega_1 < \cdots < \omega_m \le \pi$ are fixed (e.g., $\omega_k = k\pi/m$)
  – $m \gg n$ (common rule-of-thumb: $m = 15n$).

- The discretized formulation yields a relaxation of the original problem (it is possible to deal with the original problem directly, but the mathematics become very sophisticated).

- Let's now reformulate the discretized formulation in a more convenient form. Can we reformulate it as an LP?

- Recall that

$$\underset{x}{\text{minimize}} \; \underset{k}{\max} \left| f_k\left(x\right) \right|$$

can be rewritten as

$$\begin{array}{ll} \underset{t,x}{\text{minimize}} & t \\ \text{subject to} & \left| f_k\left(x\right) \right| \leq t \qquad \forall k \end{array}$$

and, equivalently, as the LP

$$\begin{array}{ll} \underset{t,x}{\text{minimize}} & t \\ \text{subject to} & -t \leq f_k\left(x\right) \leq t \qquad \forall k \;\; . \end{array}$$

- Answer: No!

- The Chebychev filter design problem cannot be recast as an LP.

- The reason is that the operator $|\cdot|$ does not denote absolute value but *magnitude* because the argument is complex-valued!

- Magnitude of a complex number:

$$|x| = |x_R + jx_I| = \sqrt{x_R^2 + x_I^2} = \left\| \begin{bmatrix} x_R \\ x_I \end{bmatrix} \right\|.$$

- The magnitude of a complex number is equivalent to the Euclidean norm of a two-dimensional vector.

- Therefore, the constraint

$$\left|H\left(\omega_k\right)-H_{\mathsf{des}}\left(\omega_k\right)\right|\leq t$$

cannot be rewritten as a linear inequality but as an SOC inequality:

$$\left\|\left[\begin{array}{c}\mathsf{Re}H\left(\omega_k\right)-\mathsf{Re}H_{\mathsf{des}}\left(\omega_k\right)\\ \mathsf{Im}H\left(\omega_k\right)-\mathsf{Im}H_{\mathsf{des}}\left(\omega_k\right)\end{array}\right]\right\|\leq t.$$

- The discretized Chebychev filter design formulation can be finally be written as an SOCP:

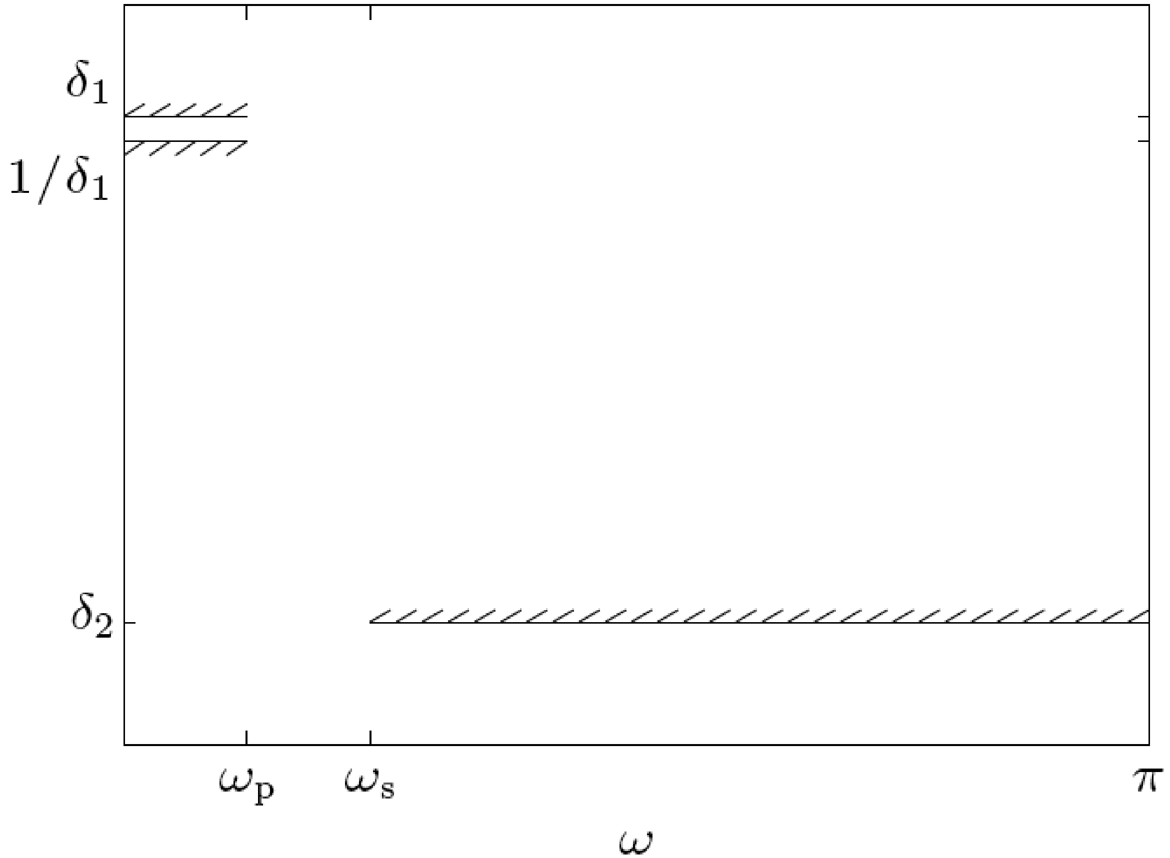
$$\begin{array}{ll} \underset{t,\mathbf{h}}{\text{minimize}} & t \\ \text{subject to} & \left\|\mathbf{A}_k\mathbf{h}-\mathbf{b}_k\right\| \leq t \qquad k=1,\cdots,m \end{array}$$

where

$$\begin{aligned} \mathbf{h} &= \left[\begin{array}{ccc} h_0 & \cdots & h_{n-1} \end{array}\right]^T \\ \mathbf{A}_k &= \left[\begin{array}{cccc} 1 & \cos\omega_k & \cdots & \cos\left(n-1\right)\omega_k \\ 0 & -\sin\omega_k & \cdots & -\sin\left(n-1\right)\omega_k \end{array}\right] \\ \mathbf{b}_k &= \left[\begin{array}{c} \mathsf{Re}H_{\mathsf{des}}\left(\omega_k\right) \\ \mathsf{Im}H_{\mathsf{des}}\left(\omega_k\right) \end{array}\right] \qquad \left(\text{note: } \mathbf{A}_k\mathbf{h} = \left[\begin{array}{c} \mathsf{Re}H\left(\omega_k\right) \\ \mathsf{Im}H\left(\omega_k\right) \end{array}\right]\right). \end{aligned}$$

Lowpass Filter Specifications

- In a lowpass filter, we have the pass frequencies in *passband* $[0, \omega_p]$ and the block frequencies in *stopband* $[\omega_s, \pi]$:



- Specifications (specs):
  - maximum passband ripple ($\pm 20 \log_{10} \delta_1$ in dB):

$$1/\delta_1 \leq |H(\omega)| \leq \delta_1, \qquad 0 \leq \omega \leq \omega_p$$

  - minimum stopband attenuation ($-20 \log_{10} \delta_2$ in dB):

$$|H(\omega)| \leq \delta_2, \qquad \omega_s \leq \omega \leq \pi.$$

- Are these nice constraints, i.e., convex?

- Recalling that the magnitude is indeed a norm,
  - the two upper-bound constraints $|H(\omega)| \leq \delta_1$ and $|H(\omega)| \leq \delta_2$ are SOC constraints
  - the lower-bound constraint $1/\delta_1 \leq |H(\omega)|$ is nonconvex!

- What can we do?

Interlude: Linear Phase Filters

- Linear phase filters satisfy:
  1. $n = 2N + 1$ is odd
  2. impulse response is symmetric about midpoint:

$$h_t = h_{n-1-t}, \qquad t = 0, \cdots, n-1.$$

- As a consequence, the frequency response can be written as

$$\begin{aligned} H(\omega) &= h_0 + h_1 e^{-j\omega} + \cdots + h_{n-1} e^{-j(n-1)\omega} \\ &= e^{-jN\omega} \left(2h_0 \cos N\omega + 2h_1 \cos (N-1)\omega + \cdots + h_N\right) \\ &\triangleq e^{-jN\omega} \tilde{H}(\omega) \end{aligned}$$

where we have used $h_0 + h_{n-1} e^{-j(n-1)\omega} = h_0 \left(1 + e^{-j2N\omega}\right) = e^{-jN\omega} h_0 2 \cos N\omega$.

- Observations on $H(\omega) = e^{-jN\omega}\tilde{H}(\omega)$ and

$$\tilde{H}(\omega) = 2h_0 \cos N\omega + 2h_1 \cos(N-1)\omega + \cdots + h_N :$$

  - term $e^{-jN\omega}$ represents an N-sample delay
  - $\tilde{H}(\omega)$ is real (this property is key)
  - same magnitude: $|H(\omega)| = |\tilde{H}(\omega)|$
  - it is called linear phase filter because the phase $\angle H(\omega)$ is linear (except for jumps of $\pm\pi$).

- How can we take advantage of these observations?

- Can we now deal with a constraint like $1/\delta_1 \leq |H(\omega)|$?

Lowpass Filter Specs

- Using $|H(\omega)| = |\tilde{H}(\omega)|$, we can rewrite the specs as

$$1/\delta_1 \leq |\tilde{H}(\omega)| \leq \delta_1, \qquad 0 \leq \omega \leq \omega_p$$

  and

$$|\tilde{H}(\omega)| \leq \delta_2, \qquad \omega_s \leq \omega \leq \pi.$$

- Noting that $|\cdot|$ now denotes absolute value instead of magnitude:
  - the two upper-bound constraints $|\tilde{H}(\omega)| \leq \delta_1$ and $|\tilde{H}(\omega)| \leq \delta_2$ are just linear constrains
  - the lower-bound constraint $1/\delta_1 \leq |\tilde{H}(\omega)|$ is still nonconvex!

- What can we do? It seems that we have not improved the problem formulation.

- Key idea:
  - the first sample at ω_1, $\tilde{H}(\omega_1)$ is either be positive or negative
  - we can assume w.l.o.g. that it is positive $\tilde{H}(\omega_1) > 1/\delta_1$ (if it's negative, use $-\mathbf{h}$ instead)
  - therefore, $|\tilde{H}(\omega_1)| = \tilde{H}(\omega_1)$
  - what about the second sample at ω_2?
  - since $\tilde{H}(\omega)$ is smooth in ω, $\tilde{H}(\omega_2)$ cannot possibly be negative, so $|\tilde{H}(\omega_2)| = \tilde{H}(\omega_2)$
  - same argument holds for all samples in the passband $\omega_k \in [0, \omega_p]$.

- As a consequence, w.l.o.g., we can substitute the nonconvex inequality $1/\delta_1 \leq |\tilde{H}(\omega)|$ by a simple linear inequality

$$1/\delta_1 \leq \tilde{H}(\omega).$$

Linear Phase Lowpass Filter Design

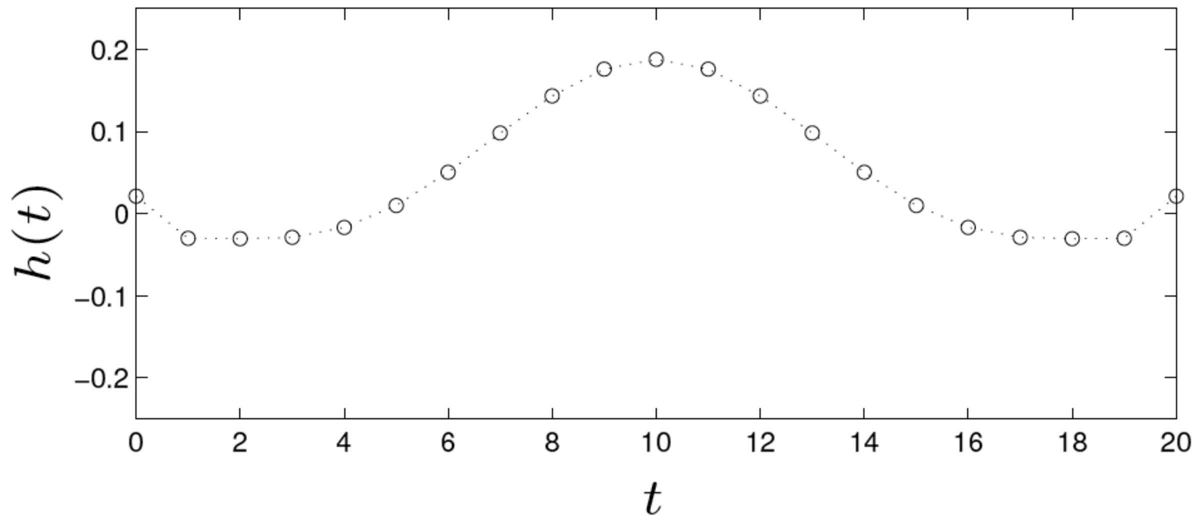
- Problem formulation for maximum stopband attenuation:

$$\begin{array}{lll} \underset{\delta_2, \mathbf{h}}{\text{minimize}} & \delta_2 & \\ \text{subject to} & 1/\delta_1 \leq \tilde{H}(\omega) \leq \delta_1, & 0 \leq \omega \leq \omega_p \\ & -\delta_2 \leq \tilde{H}(\omega) \leq \delta_2, & \omega_s \leq \omega \leq \pi. \end{array}$$

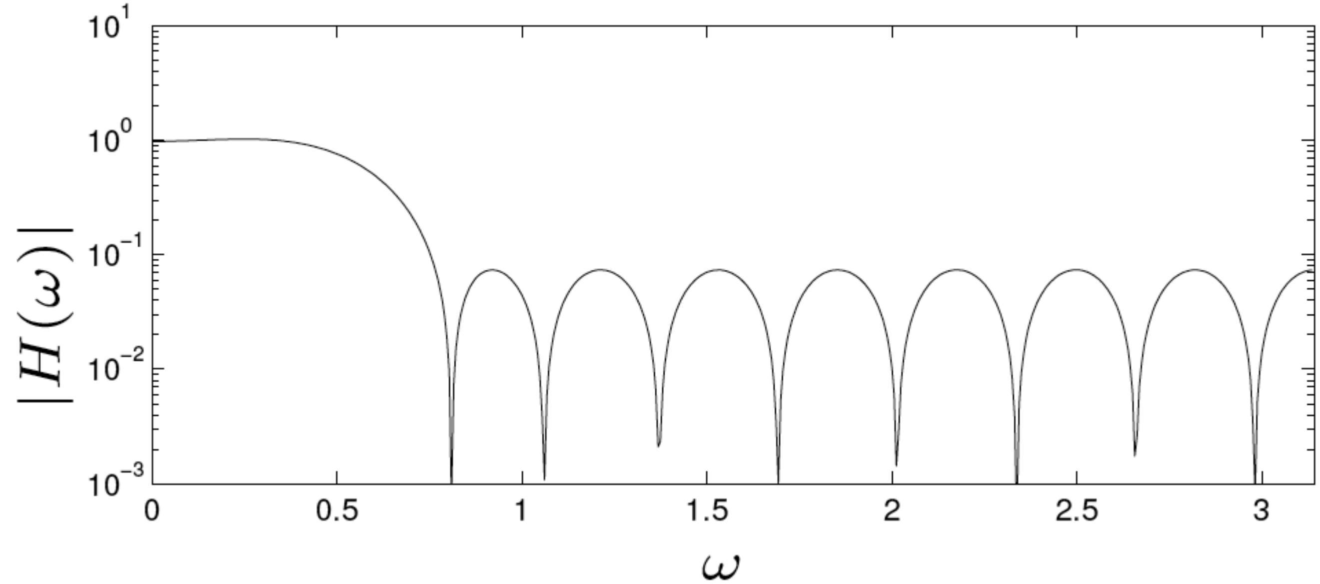
- Comments:
  - passband ripple δ_1 is given
  - the problem is an LP in variables $\delta_2, \mathbf{h}$
  - known (and used) since 1960s
  - we can add other constraints, e.g., $|h_i| \leq \alpha$.

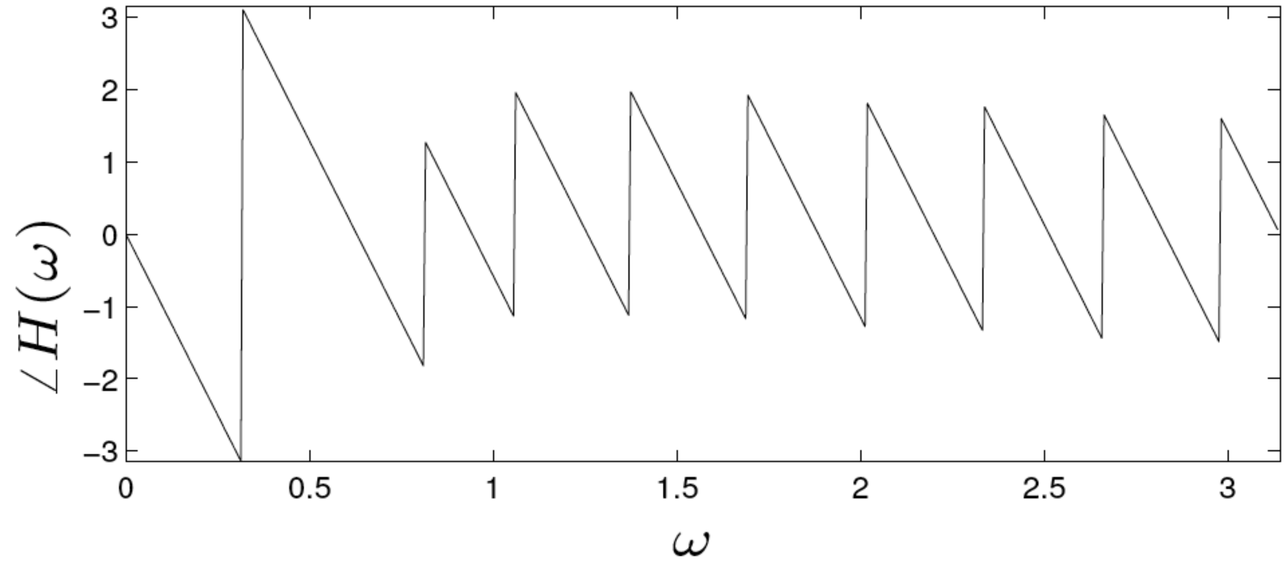
- Variations and extensions:
  - fix δ_2; minimize δ_1 (convex, but not LP)
  - fix δ_1 and δ_2, minimize ω_s (quasiconvex)
  - fix δ_1 and δ_2, minimize order n (quasiconvex).
- Example of a linear phase filter: order $n = 21$, passband $[0, 0.12\pi]$, stopband $[0.24\pi, \pi]$, max ripple $\delta_1 = 1.012$ (± 0.1 dB), design for maximum stopband attenuation.

  The impulse response $\mathbf{h}$ is



with frequency response magnitude and phase, $|H(\omega)|$ and $\angle H(\omega)$:





Filter Magnitude Specifications

- Transfer function magnitude specs have the form

$$L(\omega) \leq |H(\omega)| \leq U(\omega), \qquad \omega \in [0, \pi]$$

where $L(\omega)$ and $U(\omega)$ are the given lower and upper bounds.

- Like before:
  - the upper-bound constraint $|H(\omega)| \leq U(\omega)$ is convex
  - the lower-bound constraint $L(\omega) \leq |H(\omega)|$ is nonconvex.

- Differently from the lowpass linear phase filter design, we cannot use the same trick on the lower bound.

- What can we do?

Interlude: Autocorrelation Coefficients

- Autocorrelation coefficients are given by

$$r_t = \sum_{\tau} h_\tau h_{\tau+t}$$

(we define $h_k = 0$ for $k < 0$ or k$\geq n$).

- Some properties:
  - symmetry: $r_t = r_{-t}$
  - $r_t = 0$ for $|t| > n$
  - it suffices to specify $\mathbf{r} = [r_0, \cdots r_{n-1}]^T$.

- The Fourier transform of the autocorrelation coefficients is

$$R(\omega) = \sum_{\tau} e^{-j\omega\tau} r_\tau = r_0 + \sum_{t=1}^{n-1} 2r_t \cos \omega t = |H(\omega)|^2.$$

- Observations:
  - $R(\omega) \geq 0$ for all ω
  - $R(\omega)$ is convex in $\mathbf{h}$
  - $R(\omega)$ is linear in $\mathbf{r}$.

- How can we take advantage of the autocorrelation coefficients $\mathbf{r}$ and its Fourier transform $R(\omega)$?

- First of all, note that we can express the magnitude specifications as

$$L(\omega)^2 \leq R(\omega) \leq U(\omega)^2, \qquad \omega \in [0, \pi]$$

which are convex in $\mathbf{r}$ (in fact, linear).

- But, how does this help? Our optimization variable is $\mathbf{h}$ not $\mathbf{r}$.

- We need to reformulate the optimization problem in terms of the new optimization variable $\mathbf{r}$.

- However, once we find the optimal $\mathbf{r}$, how do we obtain the corresponding optimal $\mathbf{h}$?

- All these questions are answered by the *spectral factorization theorem*.

Spectral Factorization

- **Question**: when is $\mathbf{r}$ the autocorrelation of some $\mathbf{h}$?
- **Answer** (spectral factorization theorem): if and only if $R(\omega) \geq 0$ for all ω.
- The spectral factorization condition is convex in $\mathbf{r}$.
- The idea is then to formulate the problem using $\mathbf{r}$ as a variable (instead of $\mathbf{h}$) including the constraint $R(\omega) \geq 0$ for all ω.
- Once the problem has been solved, we know that there exists some $\mathbf{h}$ with such an autocorrelation; in fact, there are many algorithms for spectral factorization.

Log-Chebychev Magnitude-Spec Design

- In many applications it is more meaningful to work with the magnitude of the frequency response in dB instead of linear scale.

- We can then reformulate the first Chebychev problem formulation we considered

$$\text{minimize} \max_{\omega \in [0,\pi]} \left| H(\omega) - H_{\text{des}}(\omega) \right|$$

but in magnitude-dB:

$$\text{minimize} \max_{\omega \in [0,\pi]} \left| 20 \log_{10} |H(\omega)| - 20 \log_{10} D(\omega) \right|$$

where $D(\omega)$ denotes the desired frequency response magnitude ($D(\omega) > 0$ for all ω).

- We can use the spectral factorization theorem to rewrite the problem in terms of $\mathbf{r}$:

$$\begin{array}{ll} \underset{t,\mathbf{r}}{\text{minimize}} & t \\ \text{subject to} & \left|10\log_{10} R(\omega) - 10\log_{10} D^2(\omega)\right| \leq t, \qquad 0 \leq \omega \leq \pi \end{array}$$

which is still nonconvex.

- Expanding the absolute value we obtain

$$\begin{array}{ll} \underset{t,\mathbf{r}}{\text{minimize}} & t \\ \text{subject to} & -t \leq \log_{10} R(\omega) / D^2(\omega) \leq t, \qquad 0 \leq \omega \leq \pi \end{array}$$

which still is nonconvex.

- What can we do now?

- We can exponentiate the constraint:

$$\begin{array}{ll} \underset{t,\mathbf{r}}{\text{minimize}} & t \\ \text{subject to} & 10^{-t} \leq R(\omega)/D^2(\omega) \leq 10^t, \qquad 0 \leq \omega \leq \pi. \end{array}$$

- Now, define $\tilde{t} = 10^t$ and rewrite the problem finally in convex form as

$$\begin{array}{ll} \underset{\tilde{t},\mathbf{r}}{\text{minimize}} & \tilde{t} \\ \text{subject to} & 1/\tilde{t} \leq R(\omega)/D^2(\omega) \leq \tilde{t}, \qquad 0 \leq \omega \leq \pi. \end{array}$$

- Note that the spectral factorization condition is already included.

- Let's rearrange terms now. Note that $1/t \leq R(\omega)/D^2(\omega)$ can be rewritten as $D^2(\omega)/R(\omega) \leq t$.

- So we can rewrite our problem as

$$\begin{array}{ll} \underset{t,\mathbf{r}}{\text{minimize}} & t \\ \text{subject to} & \max\left\{R(\omega)/D^2(\omega), D^2(\omega)/R(\omega)\right\} \leq t, \qquad 0 \leq \omega \leq \pi. \end{array}$$

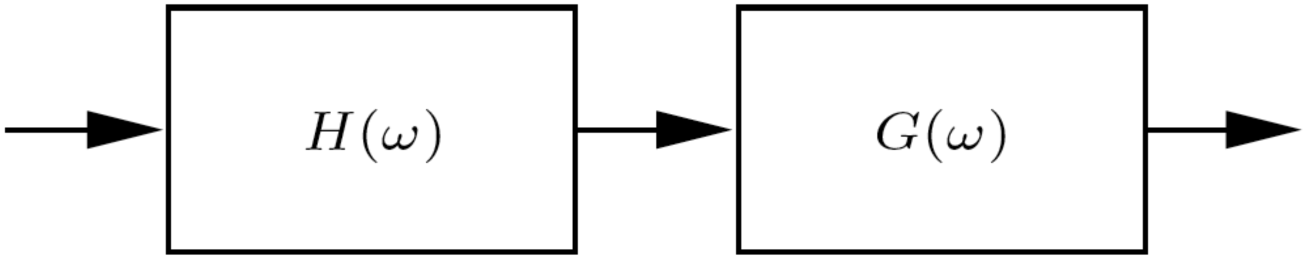
- More compactly, by defining the function $\phi(x) = \max\{x, 1/x\}$:

$$\underset{\mathbf{r}}{\text{minimize}} \ \max_{\omega \in [0,\pi]} \phi\left(R(\omega)/D^2(\omega)\right).$$

- Does this ring any bell?

Equalizer Design

- System model: concatenation of a filter $H(\omega)$, to be designed, and the unequalized channel response $G(\omega)$:



- Equalization problem: design the filter $H(\omega)$ (FIR equalizer) so that the overall response is close to the desired one $G_{\text{des}}(\omega)$:

$$H(\omega)\, G(\omega) \approx G_{\text{des}}(\omega)\,.$$

- One common choice for the desired response is $G_{\text{des}}(\omega) = e^{-jD\omega}$ (delay of D samples), i.e., equalization is deconvolution (up to a delay).

- We can add constraints on the filter coefficients $\mathbf{h}$ and $H(\omega)$ such as limits on $|h_i|$ or $\max_{\omega} |H(\omega)|$.

- A simple formulation is the **Chebychev equalizer design**:

$$\text{minimize} \max_{\omega \in [0,\pi]} |H(\omega) G(\omega) - G_{\text{des}}(\omega)|$$

  which is convex and can be reformulated as an SOCP after sampling the frequency.

- In the context of equalization, it is sometimes common to use the time domain instead of the frequency domain.

- For example, the time-domain desired response corresponding to $G_{\text{des}}(\omega) = e^{-jD\omega}$ is

$$g_{\text{des}}(t) = \begin{cases} 1 & t = D \\ 0 & t \neq D. \end{cases}$$

- Let $\tilde{g}(t)$ denote the time-domain signal corresponding to the equalized system $\tilde{G}(\omega) = H(\omega) G(\omega)$.

- **Time-domain equalization**: Inspired by the expression of $g_{\text{des}}(t)$ above, we can then formulate the filter design problem in the time domain as:

$$\begin{array}{ll} \underset{\mathbf{h}}{\text{minimize}} & \max_{t \neq D} |\tilde{g}(t)| \\ \text{subject to} & \tilde{g}(D) = 1 \end{array}$$

which is an LP.

- Variations: we can use $\sum_{t \neq D} \tilde{g}(t)^2$ or $\sum_{t \neq D} |\tilde{g}(t)|$ as objectives.

- Extensions:
  - we can impose additional convex constraints
  - we can mix the time- and frequency-domain specs
  - we can equalize multiple systems, i.e., to choose

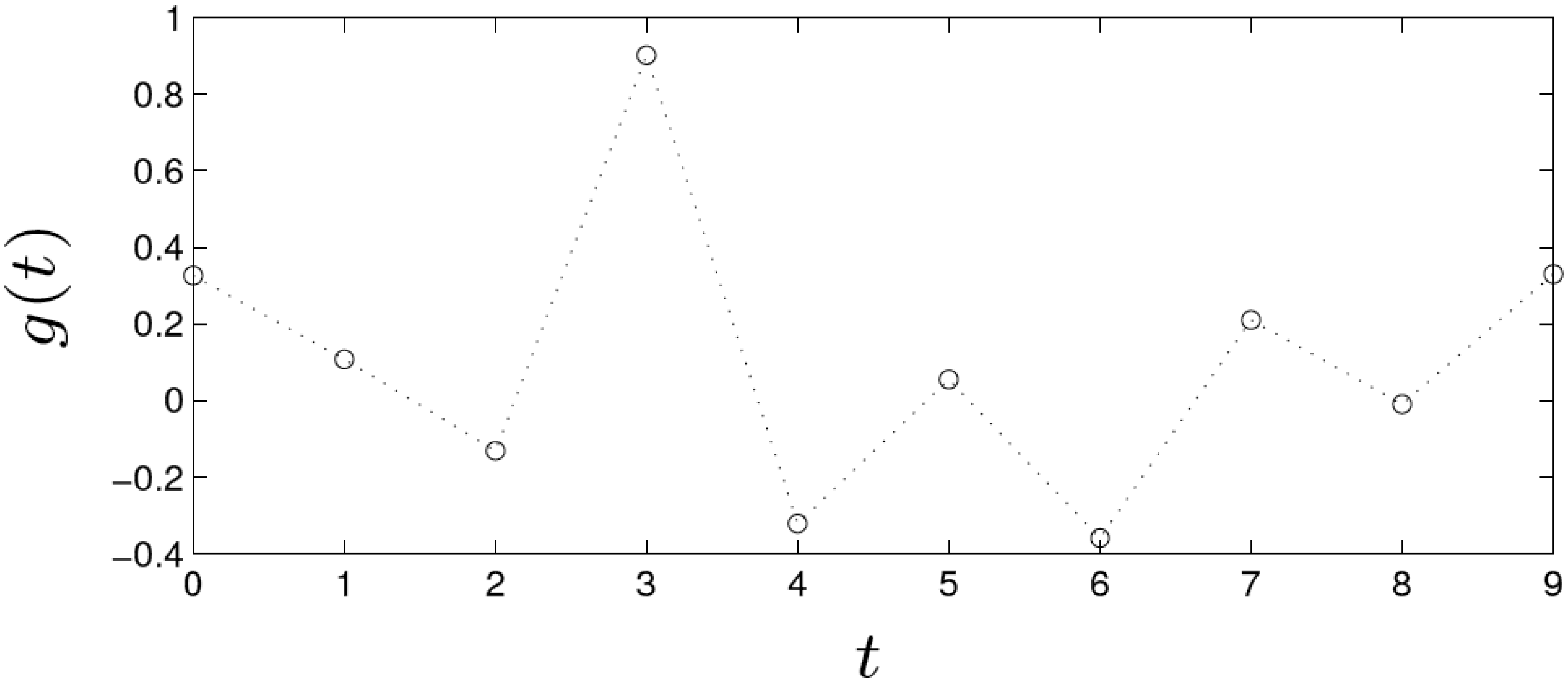
$$H(\omega) G^{(k)}(\omega) \approx G_{\mathsf{des}}(\omega), \qquad k = 1, \cdots, K$$

  - we can even equalize multi-input multi-output systems where $H(\omega)$ and $G(\omega)$ are matrices
  - it extends to multidimensional systems such as image processing.

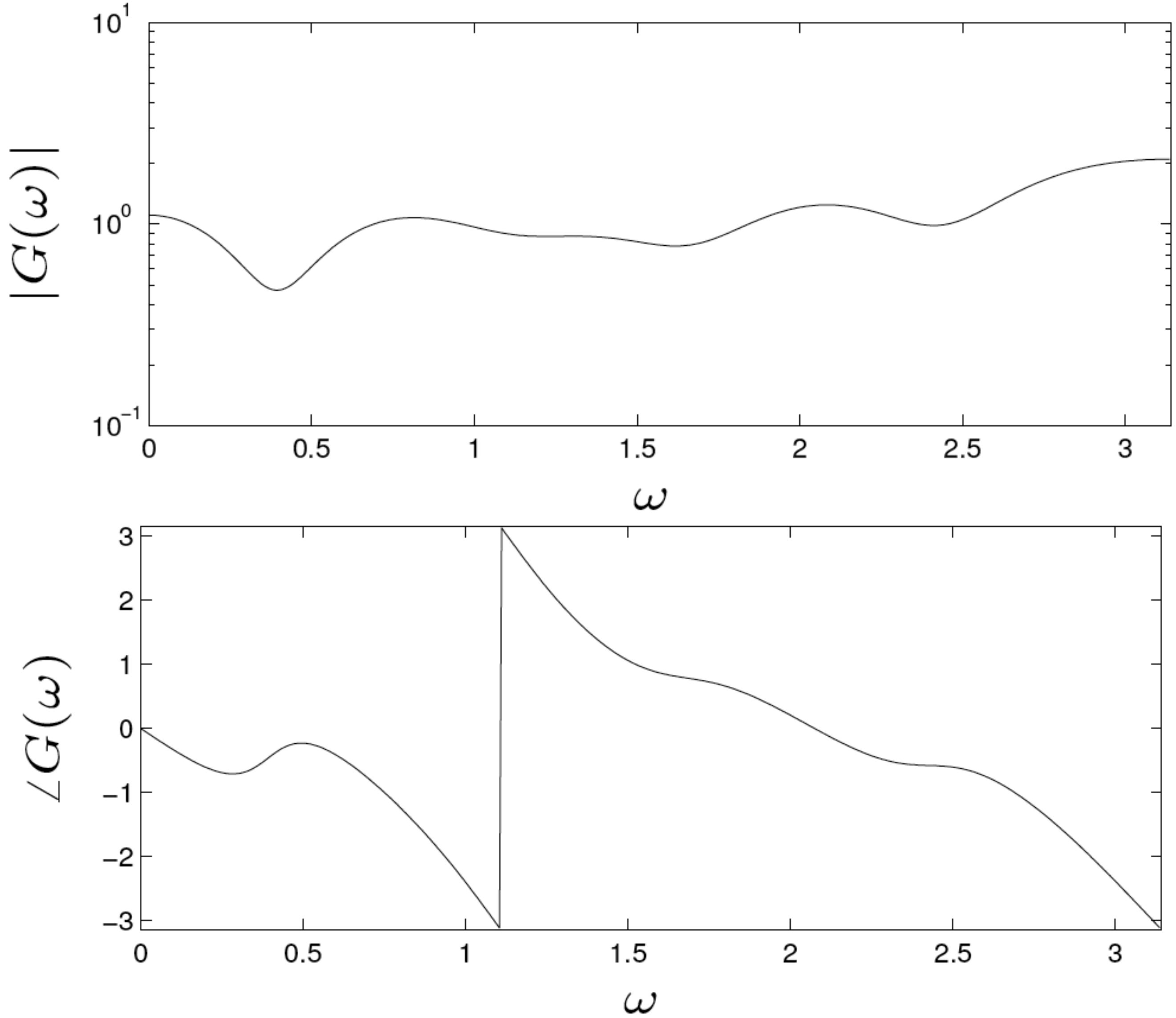
Example Filter Design

- The problem is to design a 30th order FIR equalizer with $G_{\text{des}}(\omega) = e^{-j10\omega}$.

- Consider the unequalized system $g(t)$ (10th order FIR):



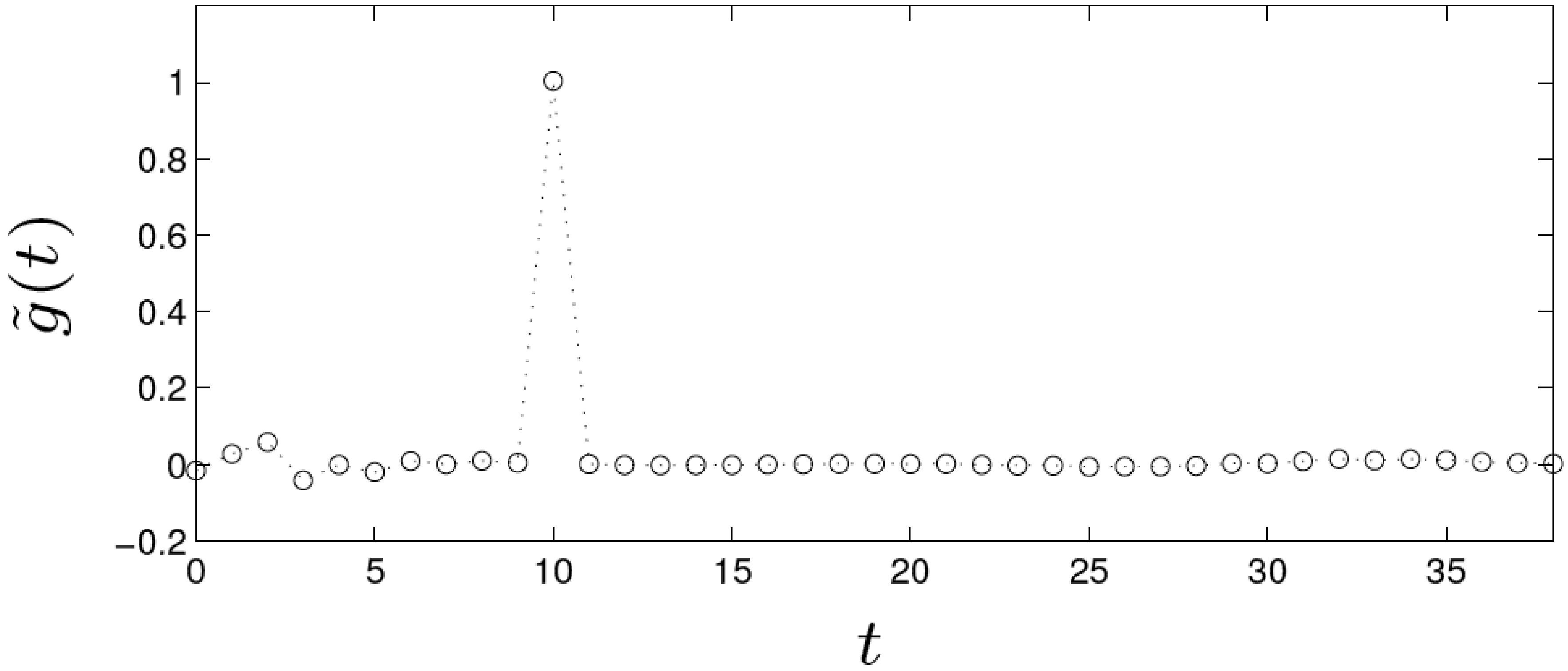
with frequency response magnitude and phase:



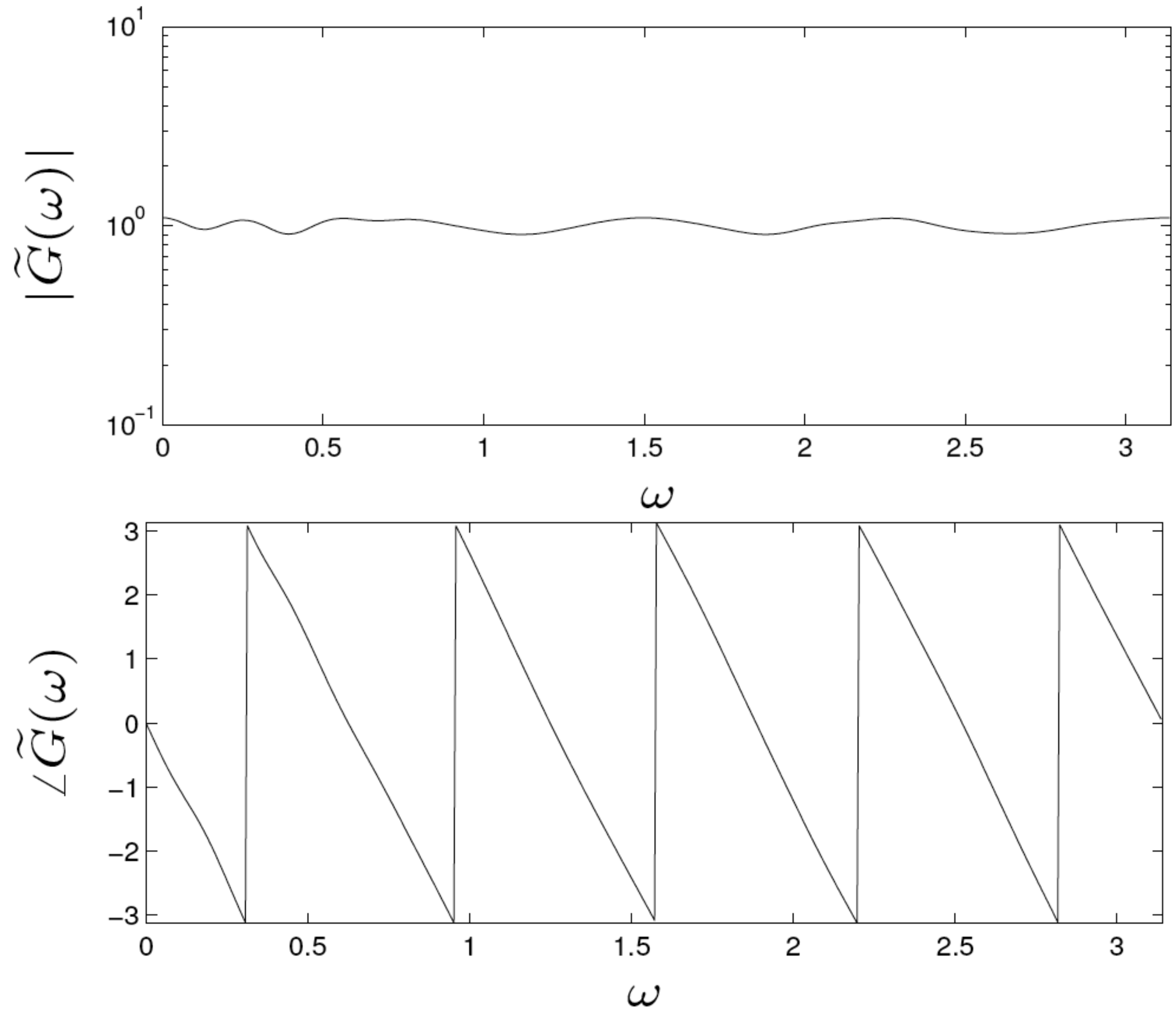
- **Chebychev equalizer design**:

$$\text{minimize} \max_{\omega \in [0,\pi]} \left| H(\omega) G(\omega) - e^{-j10\omega} \right|$$

- The equalized system impulse response $\tilde{g}(t)$ is



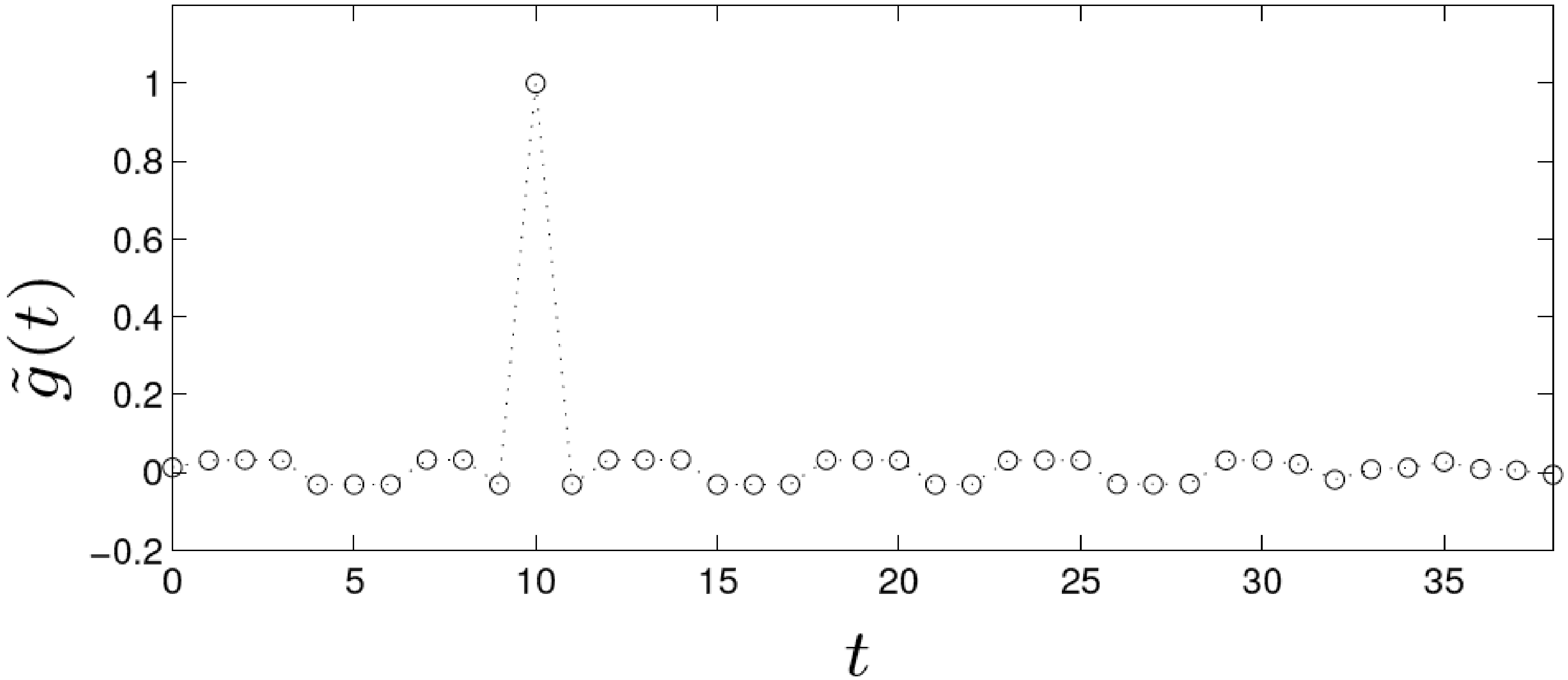
with equalized frequency response magnitude and phase:



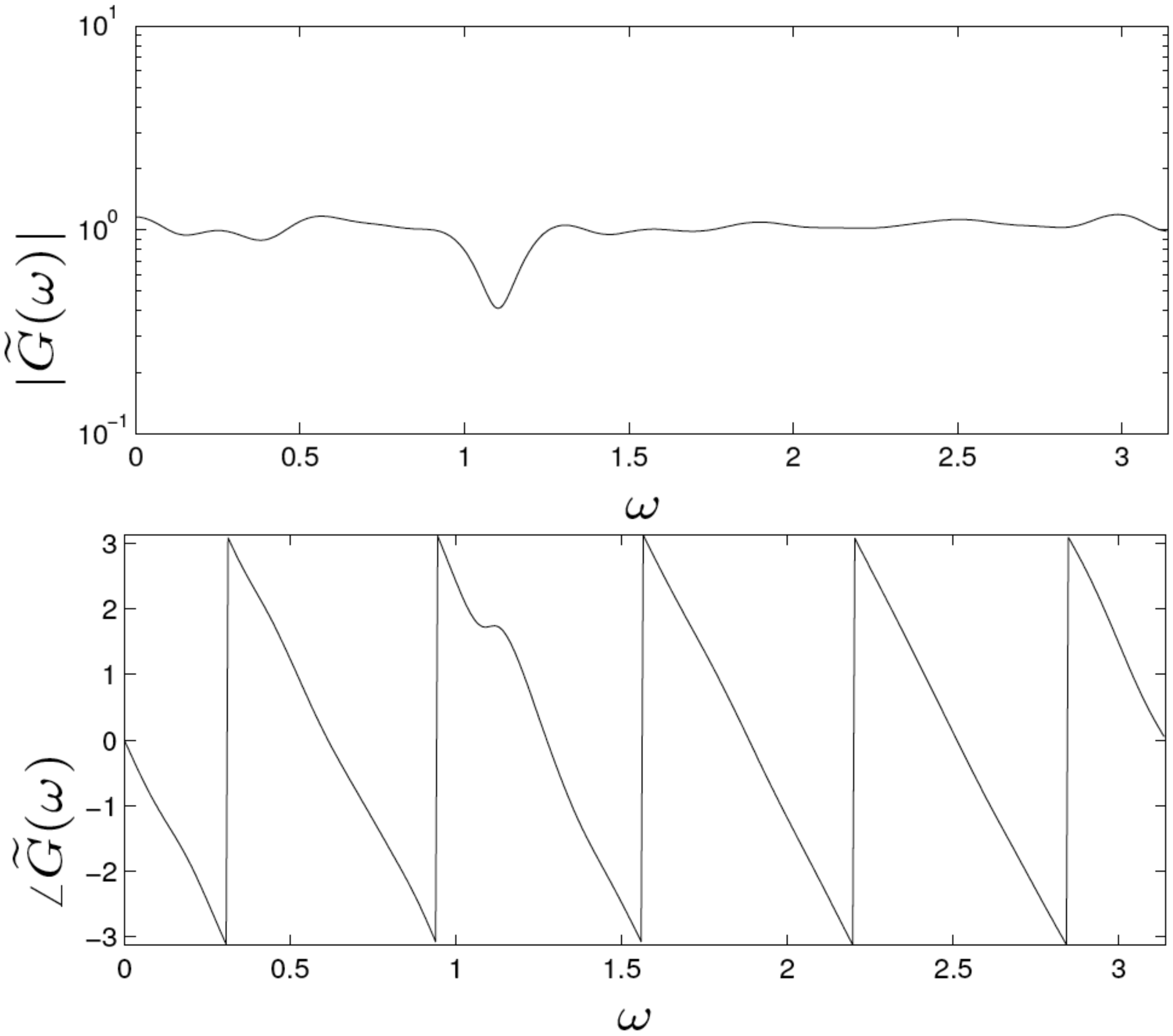
- **Time-domain equalizer design**:

$$\text{minimize} \quad \max_{t \neq 10} |\tilde{g}(t)|$$

- The equalized system impulse response $\tilde{g}(t)$ is



with equalized frequency response magnitude and phase:



Summary

- We have considered many different problem formulations of filter design:
  - Chebychev design
  - lowpass filter design
  - filter magnitude specification design
  - log-Chebychev magnitude specification design
  - equalizer design.

- Most of the formulations are initially very hard nonconvex problems.

- Using different tricks they can finally be reformulated in convex form and solved optimally.