

Capítulo 5

Projeto de Filtros Digitais

5.1 Introdução

Os filtros pertencem a uma classe muito importante de sistemas lineares invariantes no tempo, cuja função primária é selecionar, com pouca ou nenhuma atenuação, determinadas componentes de frequência e rejeitar ou remover todas as outras componentes de um sinal aplicado em sua entrada. Tais filtros são chamados de filtros seletivos em frequência. Eles são utilizados em aplicações tais como: redução de ruído, enriquecimento de sinais, equalizadores gráficos em sistemas de áudio, e em muitas outras aplicações de engenharia elétrica. Em um contexto mais amplo define-se um filtro como um dispositivo que modifica as componentes de frequência de um sinal aplicado em sua entrada.

Existem duas classes principais de filtros, os analógicos e os digitais. Os filtros analógicos utilizam componentes eletrônicos tais como resistores, capacitores, indutores e amplificadores operacionais para se construí-lo. Os filtros digitais utilizam um processador digital para realizar cálculos numéricos relativos à filtragem em um sinal amostrado. Sendo assim, além de incorporarem as vantagens do processamento digital, eles são facilmente projetados e implementados.

Neste capítulo serão descritas diversas técnicas de projeto dos filtros digitais, tanto para os filtros IIR (filtros com resposta ao impulso infinita) bem como para os FIR (filtros com resposta ao impulso finita). As técnicas, aqui apresentadas, são baseadas nos projetos clássicos de filtros analógicos, ou então na aproximação das suas características ideais. As características desejadas são especificadas pelas respostas de amplitude e de fase ou então, pela função de transferência e o projeto propriamente dito consiste em encontrar os coeficientes de um filtro IIR ou FIR que melhor aproximam as especificações desejadas.

Na prática os filtros FIR são empregados quando as especificações de projeto requerem um filtro com fase linear. Quando este requerimento não é necessário podem-se utilizar os filtros IIR. Os filtros IIR são preferidos em relação aos FIR, pois apresentam menores lóbulos laterais na banda de atenuação, e as características desejadas são obtidas com um número menor de parâmetros, reduzindo a complexidade computacional. Os filtros FIR, ao contrário, apresentam características completamente opostas aos IIR.

Neste capítulo, para facilidade de notação, representaremos a frequência angular analógica, dada em radianos por segundo, pela letra grega omega maiúscula (Ω) e a frequência em ciclos por segundo ou Hertz por (F). As frequências digitais serão representadas pelas letras minúsculas romanas w e f , respectivamente.

5.2 Projeto de filtros digitais com resposta ao impulso infinita - IIR

Os filtros IIR são filtros recursivos, cuja saída no instante atual depende dos valores presente e passados da entrada e também dos valores passados da própria saída. Eles apresentam uma resposta ao impulso com duração infinita, e como consequência eles podem ser adaptados aos filtros analógicos que também, geralmente, apresentam resposta ao impulso com duração infinita. Assim, as técnicas principais de projeto dos filtros IIR consistem na obtenção de um filtro analógico, com determinadas especificações de projeto, em seguida, na transformação deste filtro em um digital. A transformação tenta preservar alguma característica ou representação do filtro analógico.

Existem três técnicas básicas para o projeto de filtros digitais com resposta ao impulso infinita: aproximação das derivadas da equação diferencial, invariância na resposta ao impulso e transformação bilinear, que é a mais popular. Estas técnicas são baseadas na transformação de um filtro de tempo contínuo em um filtro de tempo discreto, através de um mapeamento do plano s da transformada de Laplace para o plano z . A razão para este procedimento é que os projetos clássicos dos filtros analógicos

são bem conhecidos, padronizados e tabelados, além disso, famílias de filtros padrões, como por exemplo, os filtros de Butterworth, de Chebyshev e elípticos conduzem a projetos eficientes. Uma outra razão é que estes projetos conduzem a fórmulas muito simples de transformação do domínio do tempo contínuo para o discreto.

Um filtro analógico contínuo no tempo pode ser representado, no domínio da frequência, pela seguinte função do sistema:

$$H_a(s) = \frac{\sum_{k=0}^M \beta_k s^k}{\sum_{k=0}^N \alpha_k s^k} \quad (5.1)$$

em que: α_k e β_k são os coeficientes constantes do filtro, N é ordem do filtro, que corresponde ao maior grau do denominador, e s o operador de Laplace. O subscrito 'a' indica que o filtro é analógico.

A transformada de Laplace inversa da função do sistema fornece uma outra forma de representação de um filtro, isto é, a representação através da sua resposta ao impulso.

$$h_a(t) = \frac{1}{2\pi j} \int_{\sigma-j\infty}^{\sigma+j\infty} H_a(s) e^{st} ds \quad (5.2)$$

Uma outra forma alternativa para representar um filtro analógico é descrevê-lo através de uma equação linear diferencial com coeficientes constantes como mostrado abaixo:

$$\sum_{k=0}^N \alpha_k \frac{d^k}{dt^k} y(t) = \sum_{k=0}^M \beta_k \frac{d^k}{dt^k} x(t) \quad (5.3)$$

As três representações mostradas acima conduzem a três métodos diferentes de projeto de filtros digitais IIR, como será estudado em detalhes a seguir.

5.2.1 Método por aproximação das derivadas

Este procedimento consiste em aproximar a representação por equação diferencial linear dada pela equação (5.3), por uma equação de diferenças linear. Admitindo T_a o intervalo de amostragem, uma aproximação numérica para a derivada de primeira ordem é dada pela seguinte equação:

$$\left. \frac{d}{dt} y(t) \right|_{t=nT_a} = \frac{1}{T_a} [y(n) - y(n-1)] \quad (5.4)$$

A derivada de primeira ordem apresenta como função do sistema a seguinte transformada:

$$H(s) = s$$

A equação de diferenças da equação (5.4) tem como função do sistema a seguinte transformada z:

$$H(z) = \frac{1}{T_a} [1 - z^{-1}]$$

Então, o seguinte mapeamento do plano complexo s para o plano complexo z é realizado quando se faz a seguinte aproximação numérica:

$$s = \frac{1}{T_a} [1 - z^{-1}] \quad (5.5)$$

Para a derivada de segunda ordem pode-se fazer a seguinte aproximação:

$$\left. \frac{d^2}{dt^2} y(t) \right|_{t=nT_a} = \frac{1}{T_a} \left\{ \frac{y(n) - y(n-1)}{T_a} - \frac{y(n-1) - y(n-2)}{T_a} \right\} = \frac{1}{T_a^2} \{y(n) - 2y(n-1) + y(n-2)\}$$

e a equivalência no domínio da frequência será dada por:

$$s^2 = \left[\frac{1 - z^{-1}}{T_a} \right]^2 \quad (5.6)$$

Seguindo o procedimento acima, segue que a substituição para a derivada de k-ésima ordem será:

$$s^k = \left[\frac{1 - z^{-1}}{T_a} \right]^k \quad (5.7)$$

Portanto, a função de transferência para um filtro digital IIR que aproxima a função de transferência do filtro analógico através da aproximação numérica das derivadas é realizada através da seguinte transformação:

$$H(z) = H_a(s) \Big|_{s = \frac{1}{T_a}(1 - z^{-1})} \quad (5.8)$$

As consequências desta transformação podem ser verificadas reescrevendo a equação (5.5) como segue:

$$z = \frac{1}{1 - sT_a} = \frac{1}{1 - j\Omega T_a} \Rightarrow |z| = \frac{1}{\sqrt{1 + (\Omega T_a)^2}} \quad (5.9)$$

Conforme Ω varia de $-\infty$ a ∞ no plano s, a variável z é mapeada em um círculo de raio $\frac{1}{2}$, centrado no eixo real, em $z = \frac{1}{2}$. Qualquer ponto no semiplano esquerdo do plano s, onde o filtro analógico é estável, é mapeado, no plano z, dentro do círculo de raio $\frac{1}{2}$, como mostra a figura 5.1.

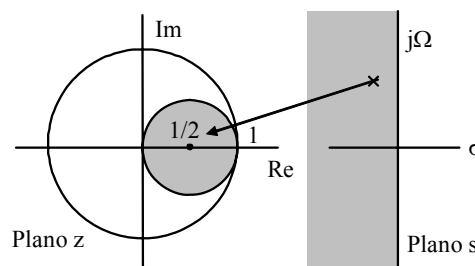


Figura 5.1: Mapeamento de s no plano z.

Consequentemente, se $H(s)$ for estável, esta transformação conduz a um filtro digital estável, pois todos os polos de $H(z)$ estarão dentro do círculo de raio unitário, mais precisamente, dentro da circunferência hachurada na figura 5.1. Observe também que a localização dos polos de $H(z)$ está confinada em uma área de frequências digitais baixas, de 0 a $\pm \pi/2$. Este fato, limita a utilização destes filtros para as características passa-baixas e passa-banda, com frequência de corte ou de ressonância baixa quando comparada com a frequência de amostragem. Não é possível projetar filtros passa-altas com esta técnica.

Exemplo 1: Converta o filtro passa-baixas com frequência de corte igual a 1 rad/s e com função de transferência mostrada abaixo, em um filtro digital IIR utilizando o método de aproximação das derivadas.

$$H_a(s) = \frac{1}{s^2 + \sqrt{2}s + 1}$$

– Para a determinação de $H(z)$ substitui-se, na equação acima, s por $(1 - z^{-1})F_a$. Assim:

$$H(z) = \frac{1/K}{1 - \frac{F_a(2F_a + \sqrt{2})}{K}z^{-1} + \frac{F_a^2}{K}z^{-2}} \quad \text{em que: } K = F_a^2 + \sqrt{2}F_a + 1$$

– Assim, os coeficientes do filtro digital em função de T_a , serão dados por:

$$b_0 = 1/K \quad a_0 = 1 \quad a_1 = -\frac{F_a(2F_a + \sqrt{2})}{K} \quad a_2 = \frac{F_a^2}{K}$$

Na figura 5.2 são mostrados espectros de amplitude dos filtros para frequências de amostragem de 2 Hz e 10 Hz. Observe que uma melhor aproximação para o filtro analógico ocorre para a frequência de amostragem maior ($F_a = 10\text{Hz}$). Para frequências de amostragem grandes, os intervalos de amostragem são pequenos e as frequências digitais de interesse ficam concentradas em torno do eixo real, onde os círculos de raio $\frac{1}{2}$ e unitário se tocam, assim neste intervalo melhor será a aproximação.

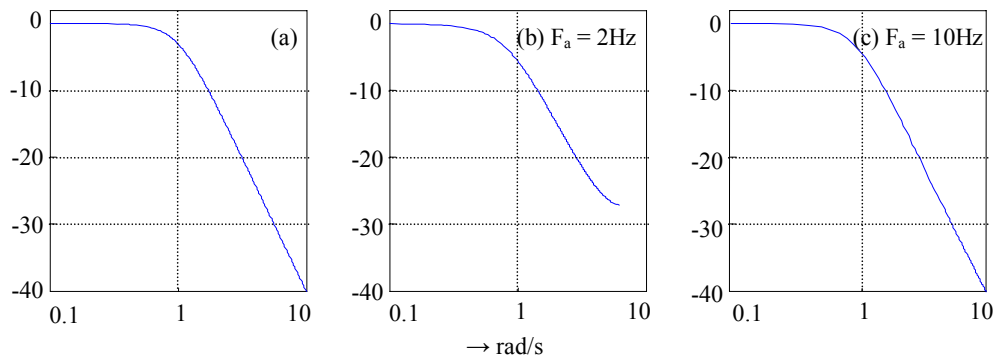


Figura 5.2: Projeto de filtros por aproximação das derivadas: (a) analógico; (b) digital; (c) digital.

5.2.2 Método por invariância da resposta ao impulso

Este método tem como objetivo a obtenção de um filtro IIR cuja resposta à amostra unitária, $h(n)$, seja uma versão amostrada da resposta ao impulso do filtro analógico que se quer aproximar. Para preservar a resposta ao impulso, $h_a(t)$ é amostrada em intervalos regulares T_a de modo a se obter a resposta $h(n)$ do filtro digital tal que:

$$h(n) = h_a(nT_a) = h_a(t)|_{t=nT_a} \quad (5.10)$$

Lembrando que a equação no domínio da frequência, que relaciona o espectro de frequências de um sinal amostrado a uma taxa F_a com o espectro do sinal analógico é dada por:

$$H(f) = F_a \sum_{k=-\infty}^{\infty} H_a[(f - k)F_a], \quad (5.11)$$

fica claro, pelo critério de Nyquist, que a resposta em frequência do filtro digital se aproxima do analógico, desde que a frequência de amostragem seja estabelecida suficientemente grande, de modo a minimizar o efeito da superposição espectral (*aliasing*). Conseqüentemente, este método é também limitado para os filtros passa-baixas e passa-banda, nos quais não existem oscilações na banda de atenuação. Ele é inadequado para o projeto de filtros passa-altas e rejeita banda a não ser que um filtro *antialiasing* seja utilizado.

Para mostrar o desenvolvimento do projeto de filtros por invariância da resposta ao impulso, considere um filtro analógico de ordem N , com N polos simples e distintos. Assim, a função de transferência na forma de frações parciais é dada por:

$$H_a(s) = \sum_{k=1}^N \frac{A_k}{s - p_k} \quad (5.12)$$

em que: p_k são os polos do filtro analógico e A_k são os coeficientes da expansão em frações parciais.

A resposta ao impulso é determinada pela transformada de Laplace inversa da equação (5.12):

$$h_a(t) = \sum_{k=1}^N A_k e^{p_k t} u(t) \quad (5.13)$$

Para preservar o formato da resposta ao impulso, $h(t)$ é amostrada nos instantes $t = nT_a$. Assim,

$$h(n) = h_a(nT_a) = \sum_{k=1}^N A_k e^{p_k n T_a} u(n) \quad (5.14)$$

Calculando a transformada z de $h(n)$ tem-se:

$$H(z) = \sum_{n=0}^{\infty} h(n) z^{-n} = \sum_{n=0}^{\infty} \sum_{k=1}^N A_k e^{p_k n T_a} z^{-n} = \sum_{k=1}^N A_k \sum_{n=0}^{\infty} \left(e^{p_k T_a} z^{-1} \right)^n$$

portanto:

$$H(z) = \sum_{k=1}^N \frac{A_k}{1 - e^{p_k T_a} z^{-1}} \quad (5.15)$$

A função de transferência acima corresponde a um filtro digital que apresenta $h(n) = h_a(nT_a)$ desde que não ocorra superposição espectral. Note na equação (5.15) que os polos, d_k , deste filtro são dados por:

$$d_k = e^{p_k T_a} \quad : k = 1, 2, \dots, N \quad (5.16)$$

As equações (5.15) e (5.16) fornecem um meio para se projetar filtros digitais pelo método de invariância ao impulso. Note que não é necessária a determinação de $h_a(t)$ para o projeto, como se poderia pensar a princípio. Na realidade são necessários o conhecimento dos polos do protótipo analógico e a frequência de amostragem do sinal.

Com relação à estabilidade do filtro pode-se notar diretamente que se o filtro analógico de referência é estável, então os seus polos $p_k = \sigma_k + j\Omega_k$ apresentam $\sigma_k < 0$. Neste caso, como:

$$|d_k| = e^{\sigma_k} < 1 \quad (5.17)$$

Assim, este método conduz a um filtro causal, estável, pois desde que os polos de $H_a(s)$ estão localizados no semiplano esquerdo do plano s , então os polos de $H(z)$ estarão localizados dentro do círculo de raio unitário.

Exemplo 2: Converta o filtro do exemplo 1, em um filtro digital pelo método de invariância ao impulso.

$$H_a(s) = \frac{1}{s^2 + \sqrt{2}s + 1}$$

- Calculando os polos de $H_a(s)$ e expandindo em frações parciais tem-se que:

$$H_a(s) = \frac{-\frac{\sqrt{2}}{2}j}{s + \frac{\sqrt{2}}{2}(1-j)} + \frac{\frac{\sqrt{2}}{2}j}{s + \frac{\sqrt{2}}{2}(1+j)}$$

- Polos de $H_a(s)$ e de $H(z)$:

$$p_{1,2} = -\frac{\sqrt{2}}{2}(-1 \pm j) \quad \Rightarrow \quad d_{1,2} = e^{\frac{\sqrt{2}}{2}(-1 \pm j)T_a}$$

- Cálculo de $H(z)$:

$$H(z) = \frac{-\frac{\sqrt{2}}{2}j}{1 - e^{\frac{\sqrt{2}}{2}(-1+j)T_a} z^{-1}} + \frac{\frac{\sqrt{2}}{2}j}{1 - e^{\frac{\sqrt{2}}{2}(-1-j)T_a} z^{-1}}$$

$$H(z) = \frac{\sqrt{2}e^{-\frac{\sqrt{2}}{2}T_a} \sin\left(\frac{\sqrt{2}}{2}T_a\right) z^{-1}}{1 - 2e^{-\frac{\sqrt{2}}{2}T_a} \cos\left(\frac{\sqrt{2}}{2}T_a\right) z^{-1} + e^{-\sqrt{2}T_a} z^{-2}}$$

- Assim, os coeficientes do filtro digital em função de F_a serão:

$$b_0 = 0 \quad b_1 = \sqrt{2}e^{-\frac{\sqrt{2}}{2}/F_a} \sin\left(\frac{\sqrt{2}}{2}/F_a\right) \quad a_0 = 1 \quad a_1 = 2e^{-\frac{\sqrt{2}}{2}/F_a} \cos\left(\frac{\sqrt{2}}{2}/F_a\right) \quad a_2 = e^{-\sqrt{2}/F_a}$$

- Para manter o ganho compatível com o analógico multiplicamos $H(z)$ por T_a , ou seja, multiplicam-se todos os coeficientes b_k acima pelo período de amostragem.

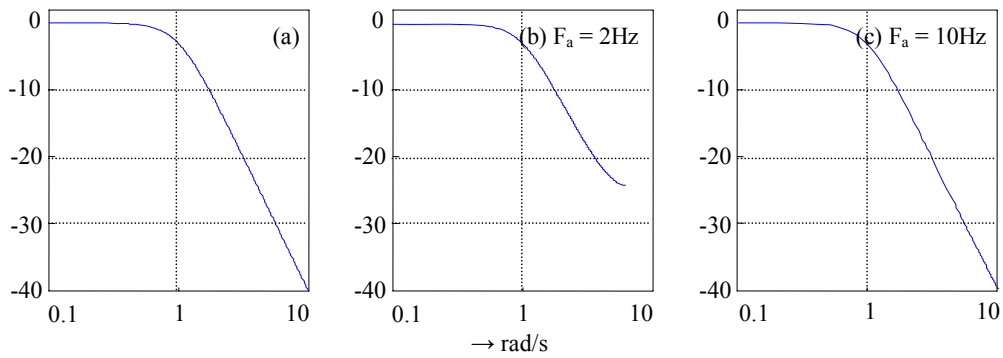


Figura 5.3: Filtro digital por invariância ao impulso: (a) analógico; (b) digital; (c) digital.

Uma vantagem na utilização deste método é que as frequências Ω e ω são relacionadas linearmente, mas como o espectro de $H(z)$ é o mesmo de $H(s)$ só que se repete em múltiplos da frequência de amostragem alguma superposição de espectros é esperada e em alguns casos pode-se tornar intolerável.

5.2.3 Método por transformação bilinear

Este método é o melhor procedimento para se projetar filtros digitais a partir de um protótipo analógico. Ele supera as limitações dos anteriores, que são restritos aos filtros passa-baixas e passa-banda. Ele evita o problema da superposição espectral, enquanto mantém as principais características da função de transferência do filtro analógico. A transformação bilinear faz o mapeamento de todo o eixo $-\infty < \Omega < \infty$ no círculo unitário $-\pi < \omega < \pi$, através da utilização da regra de integração numérica (regra do trapézio). O uso desta técnica é restrito às situações onde é aceitável distorção no eixo das frequências por causa da compressão de Ω entre $\pm\infty$ para ω entre $\pm\pi$.

Considere um sistema de primeira ordem com função de transferência:

$$H_a(s) = \frac{Y(s)}{X(s)} = \frac{b}{s+a} \quad (5.18)$$

A equação diferencial associada a este sistema é dada por:

$$\frac{d}{dt} y(t) + ay(t) = bx(t) \quad (5.19)$$

Sabe-se que a integral da derivada de $y(t)$ fornece:

$$y(t) = \int_{t_0}^t y'(t) dt + y(t_0) \quad (5.20)$$

em que $y'(t)$ é a derivada de $y(t)$.

Admitindo $t = nT_a$ e $t_0 = (n-1)T_a$ a integral é realizada em um intervalo correspondente a duas amostras consecutivas. Assim, aplicando a regra do trapézio na equação (5.20) tem-se que:

$$y(n) = \frac{T_a}{2} [y'(n) + y'(n-1)] + y(n-1) \quad (5.21)$$

Admitindo $t = nT_a$ na equação (5.19) e substituindo na equação (5.21), acima, segue que:

$$\left[1 + \frac{aT_a}{2}\right]y(n) - \left[1 - \frac{aT_a}{2}\right]y(n-1) = \frac{bT_a}{2}[x(n) + x(n-1)] \quad (5.22)$$

A equação acima é uma aproximação numérica do sistema de primeira ordem mostrado em (5.18) e em (5.19). Assim, a transformada z da equação (5.22) fornece a função do sistema digital equivalente:

$$H(z) = \frac{b}{\frac{2}{T_a} \left[\frac{1-z^{-1}}{1+z^{-1}} \right] + a} \quad (5.23)$$

Comparando as equações 5.18) e (5.23) pode-se concluir que o mapeamento do plano s para o plano z pode ser feito pela seguinte relação:

$$s = \frac{2}{T_a} \left[\frac{1-z^{-1}}{1+z^{-1}} \right] \quad (5.24)$$

Assim, substituindo s pela relação acima em H(s) obtém-se o filtro digital equivalente pelo método da transformação bilinear.

Para verificar as consequências desta transformação, seja $s = j\Omega$ e $z = e^{jw}$. Após alguma manipulação algébrica na equação (5.24) tem-se que:

$$\Omega = \frac{2}{T_a} \tan\left(\frac{w}{2}\right) \quad \text{ou} \quad w = 2 \tan^{-1}\left(\frac{\Omega T_a}{2}\right) \quad (5.25)$$

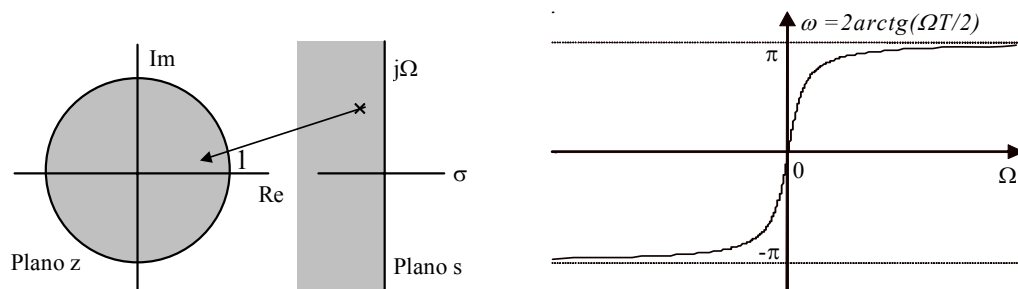


Figura 5.4: Mapeamento para a transformação bilinear.

Observe que a faixa de frequências entre 0 e ∞ , é mapeada entre 0 e π , no plano z, e a faixa $-\infty < \Omega < 0$ é mapeada no círculo unitário tal que $-\pi < w < 0$.

A transformação bilinear evita o problema da superposição espectral encontrado nos outros dois métodos, pois o eixo $j\Omega$ é mapeado inteiramente no círculo unitário do plano z, desse modo não há superposição espectral no domínio da frequência. Todo o semiplano esquerdo do plano s é mapeado dentro do círculo unitário o que indica que a transformação é estável. Contudo, este mapeamento é altamente não linear, devido à função tangente que relaciona Ω com w na equação (5.25). Esta relação comprime o eixo $j\Omega$. Este tipo de compressão é chamado na literatura de *frequency warping* e deve, a princípio, ser tolerado quando da utilização deste método.

Exemplo 3: Converta o filtro do exemplo 1, em um filtro digital utilizando o método de transformação bilinear.

$$H_a(s) = \frac{1}{s^2 + \sqrt{2}s + 1}$$

– substituindo s por $2F_a \left[\frac{1-z^{-1}}{1+z^{-1}} \right]$ tem-se que:

$$H(z) = \frac{(1 + 2z^{-1} + z^{-2})/k_1}{1 - \frac{8F_a^2 - 2}{k_1}z^{-1} + \frac{1 - 2\sqrt{2}F_a + 4F_a^2}{k_1}z^{-2}} \quad \text{em que: } k_1 = 1 + 2\sqrt{2}F_a + 4F_a^2$$

Na figura 5.4 são mostrados espectros de amplitude dos filtros para frequências de amostragem de 2 Hz e 10 Hz. Observe que há pouca diferença na aproximação do filtro analógico tanto para a frequência de amostragem de 2 Hz quanto para a de 10Hz, a não ser pela distorção no eixo de frequência. Existe uma compressão na resposta de amplitude do filtro digital. Uma das grandes vantagens na utilização deste método é não há restrição no tipo de filtro que será projetado, portanto este método é o mais utilizado na obtenção de filtros digitais IIR.

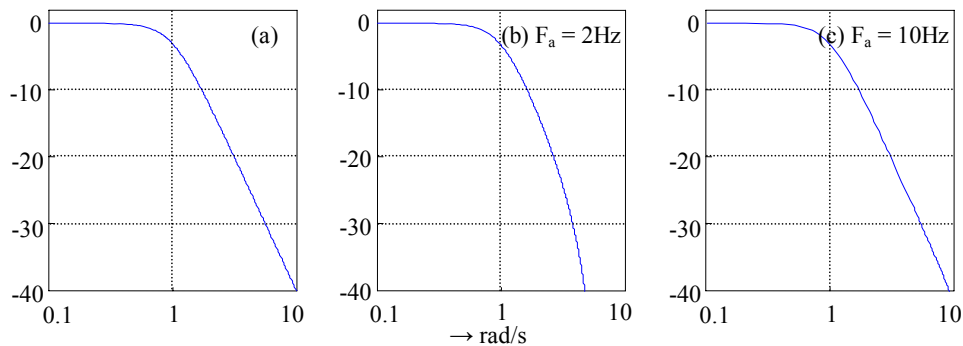


Figura 5.4: Filtro digital utilizando transformação bilinear: (a) analógico; (b) digital; (c) digital.

5.3 Família de respostas de filtros analógicos

A técnica mais comum de projeto de filtros IIR é a que transforma um protótipo analógico, previamente projetado, em um digital, utilizando um dos três métodos mostrados na seção anterior. Nesta seção serão estudados alguns procedimentos de projeto dos filtros analógicos, bem como algumas famílias de respostas de filtros analógicos muito empregadas na prática e que serão utilizadas como base nos projetos dos filtros digitais IIR. As famílias, aqui discutidas, são apresentadas para os protótipos passa-baixas, os outros tipos de filtros podem ser facilmente obtidos por transformação de frequências.

5.3.1 Resposta de Butterworth

Sabendo-se que a característica de amplitude de um filtro ideal é plana na faixa de passagem, o filtro de Butterworth tenta aproximar esta característica através de uma função polinomial no domínio da frequência. Ele é caracterizado por uma resposta de amplitude que é maximamente plana na faixa de passagem (figura 5.5.a) e por uma resposta decrescente na banda de parada. O módulo ao quadrado da sua função de transferência é definido como:

$$|H(j\Omega)|^2 = \frac{1}{1 + \left(\frac{\Omega}{\Omega_c}\right)^{2N}} = \frac{1}{1 + \left(\frac{-s^2}{\Omega_c^2}\right)^N} \quad (5.26)$$

em que: Ω_c é a frequência de corte do filtro (queda de 3 dB na banda de passagem) e N é a ordem do filtro.