

**Mestrado Integrado em Engenharia Informática e Computação**

**4º Ano**

**2º Semestre**

**Amostragem e Quantização de Sinais Media**

**Multimédia e Novos Serviços**

**EIC0064**

**2017/2018**

**Paulo Jorge Silva Ferreira,** [**up201305617@fe.up.pt**](mailto:up201305617@fe.up.pt)

**20 de fevereiro de 2018**

Índice

[Experiência 1 – Introdução 3](#_Toc506909201)

[Experiência 2 – Amostragem 4](#_Toc506909202)

[Sem Filtro 5](#_Toc506909203)

[Com Filtro 7](#_Toc506909204)

[Experiência 3 – Quantização 12](#_Toc506909205)

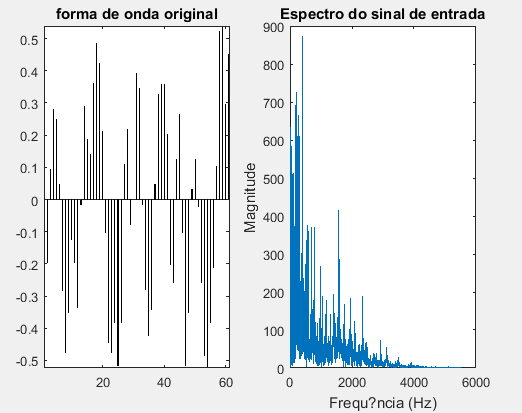
# Experiência 1 – Introdução

Foram gerados dois ficheiros áudio, um com uma taxa de amostragem de 11025*Hz* e outro com 44100*Hz*, recorrendo ao programa *VLC Media Player*. A taxa de amostragem é a quantidade de [amostras](https://pt.wikipedia.org/wiki/Amostragem_de_sinal) de um [sinal analógico](https://pt.wikipedia.org/wiki/Sinal_anal%C3%B3gico) reunidas numa determinada unidade de tempo, para conversão num  [sinal digital](https://pt.wikipedia.org/wiki/Sinal_digital). Quanto maior for a taxa de amostragem (mais amostras por unidade de tempo) mais fácil será reproduzir o sinal original, ou seja, a cópia é mais fiel.

Tendo isso em conta e depois de ouvir os dois ficheiros áudio gerados, chegou-se à conclusão que o ficheiro gerado com uma taxa de amostragem de 44100*Hz* apresenta uma maior qualidade em comparação com o de 11025*Hz*. Também é possível perceber que o som com uma maior taxa de amostragem é muito semelhante ao som original e que o com a taxa mais baixa apresenta uma qualidade bastante inferior ao original.

# Experiência 2 – Amostragem

Para realizar a experiência 2 foi utilizado o ficheiro de áudio “*batman\_theme\_x.wav*”, que se encontra na página do Moodle da Unidade Curricular. Na figura seguinte encontra-se a forma de onda original e o espetro do sinal de entrada do som utilizado.



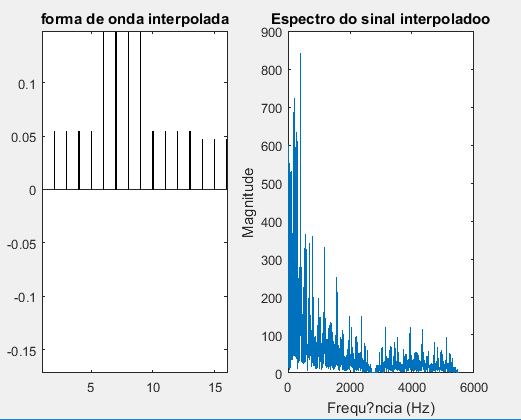
*Fig. 1 – Forma de onda original e espectro do sinal de entrada*

O ficheiro áudio de entrada utilizado é um ficheiro não comprimido com o formato *.wav* e possui uma frequência de amostragem de 44100*Hz*.

Utilizando o Teorema de *Nyquist*, , onde , pode-se concluir que, para evitar *aliasing*, a frequência máxima do sinal de entrada (frequência de *Nyquist*) deve ser menor ou igual a 22050*Hz*, sendo esta a frequência de corte ideal.

## Sem Filtro

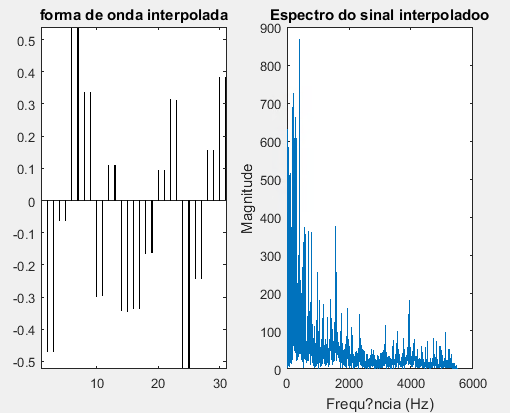
Na figura seguinte é possível observar a forma de onda interpolada e o espectro do sinal interpolado obtidos após uma amostragem de fator igual a 4 ().



*Fig. 2 – Forma de onda interpolada e espectro do sinal interpolado (k=4)*

Da análise dos espectros do sinal de entrada e do sinal interpolado, para k=4, pode-se concluir que no espectro do sinal interpolado as frequências acima da frequência de *Nyquist* aumentaram de magnitude relativamente ao espectro do sinal de entrada.

Na figura seguinte é possível observar a forma de onda interpolada e o espectro do sinal interpolado obtidos após uma amostragem de fator igual a 2 ().



*Fig. 3 – Forma de onda interpolada e espectro do sinal interpolado (k=2)*

Mais uma vez, da análise dos espectros do sinal de entrada e do sinal interpolado, para k=2, pode-se concluir que no espectro do sinal interpolado as frequências acima da frequência de *Nyquist* aumentaram de magnitude relativamente ao espectro do sinal de entrada, tal como ocorreu no caso anterior. No entanto este aumento é inferior ao verificado quando k=4, ocorrendo ainda assim alguns picos com magnitude superior.

Do ponto de vista percetual o sinal obtido com k=2 apresenta uma melhor qualidade quando comparado com o sinal obtido para k=4, mas ainda assim com qualidade inferior ao som original.

Para k=4, foi obtido um Erro Quadrático Médio (MSE) de 0.0429577 e um Peak Signal-to-Noise Ratio (PSNR) de 13.6015. Para k=2, MSE = 0.00962606 e PSNR = 20.0974.

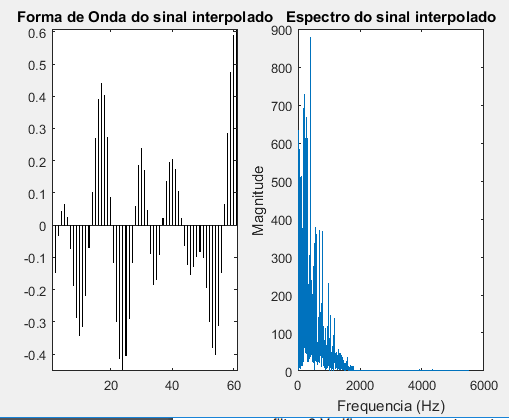
O Peak Signal-to-Noise Ratio é a relação entre a máxima energia de um sinal e o ruído que afeta a sua representação fidedigna. Quanto maior o PSNR, maior a qualidade da reconstrução.

Tendo em conta o PSNR e a perceção auditiva pode-se concluir que o sinal obtido com k=2 é de qualidade superior ao obtido com k=4.

O MSE para k=2 também é inferior quando comparado com o MSE obtido para k=4, concluindo-se que o sinal com o MSE mais baixo está mais próximo do sinal de entrada, uma vez que existem menos erros.

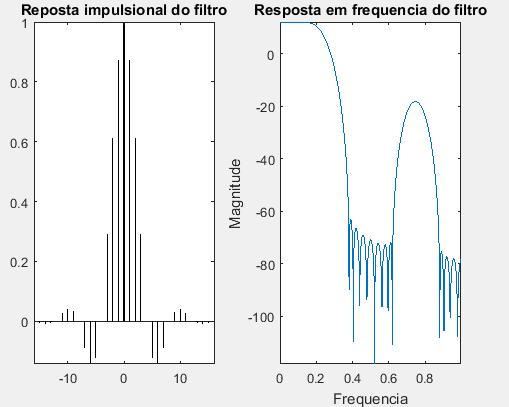
## Com Filtro

Na figura seguinte é possível observar a forma de onda interpolada e o espectro do sinal interpolado obtidos após uma amostragem de fator igual a 4 ().



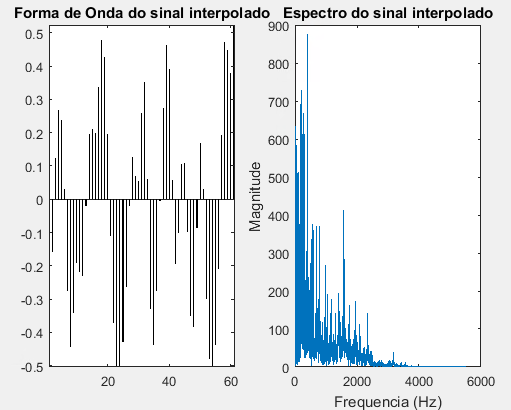
*Fig. 4 – Forma de onda e espectro do sinal interpolado com filtro (k=4)*

O espetro do sinal descrito pode ser explicado com base na resposta em frequência do uso filtro, apresentado na figura seguinte, onde se verifica um grande declive na região dos 0,22KHz (22050Hz) como se tinha sido previsto inicialmente.



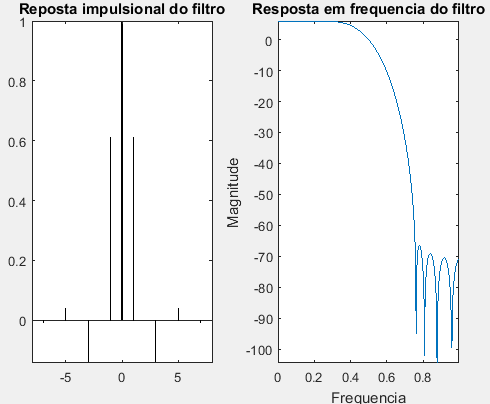
*Fig. 5 – Resposta impulsional e em frequência do filtro*

Na figura seguinte é possível observar a forma de onda interpolada e o espectro do sinal interpolado obtidos após uma amostragem de fator igual a 2 ().



*Fig. 6 – Forma de onda e espectro do sinal interpolado com filtro (k=2)*

O espetro do sinal descrito pode ser explicado com base na resposta em frequência do uso filtro, apresentado na figura seguinte, onde se verifica um grande declive na região dos 0,30KHz (30000Hz).



*Fig. 7 – Resposta impulsional e em frequência do filtro*

Apesar de ainda ser possível observar uma atenuação das frequências superiores à frequência de Nyquist, este efeito não é tão acentuado comparativamente aos resultados obtidos, nas mesmas condições, para um fator de 4.

Para k=4, foi obtido um Erro Quadrático Médio (MSE) de 0.013163 e um Peak Signal-to-Noise Ratio (PSNR) de 18.7383. Para k=2, MSE = 0.000824002 e PSNR = 30.7726.

Tendo em conta o PSNR e a perceção auditiva pode-se concluir que o sinal obtido com k=2 é de qualidade superior ao obtido com k=4.

O MSE para k=2 também é inferior quando comparado com o MSE obtido para k=4, concluindo-se que o sinal com o MSE mais baixo está mais próximo do sinal de entrada, uma vez que existem menos erros.

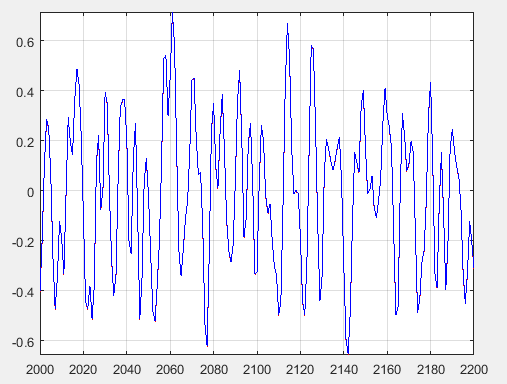
Verifica-se também valores superiores de MSE e PSNR aos obtidos nas mesmas condições, mas sem filtros, ou seja, conduziram a melhores resultados. Conclui-se assim que amostragens com filtros produzem melhores resultados, atenuando frequências indesejadas.

# Experiência 3 – Quantização

Para realizar a experiência 2 foi utilizado o ficheiro de áudio “*batman\_theme\_x.wav*”, que se encontra na página do Moodle da Unidade Curricular. Após a quantização da amostra original em 256 níveis obteve-se a forma de onda a seguir apresentada.

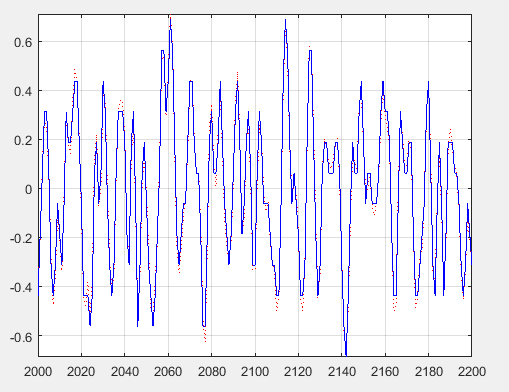
Como se pode observar a quantidade de pontos vermelhos dispostos fora das linhas azuis (desvios entre a amostra original e o resultado obtido) é reduzida, o que permite concluir que o resultado obtido após a quantização se aproxima ao original. Fato este que se confirmou após a audição do resultado final, não sendo percetíveis diferenças significativas entre a amostra original e a quantizada.

Para N=256, foi obtido um Erro Quadrático Médio (MSE) de 1.52588e-05 e um Peak Signal-to-Noise Ratio (PSNR) de 23.6994. Estes valores confirmam a perceção obtida pela audição do resultado final, uma vez que o MSE é bastante baixo, indicando poucas diferenças entre o som original e o quantizado, e o PSNR é alto e quanto maior o PSNR, maior a qualidade da reconstrução.



*Fig. 8 –* Comparação do resultado da quantização com amostra original (N = 256)

Como se pode observar a quantidade de pontos vermelhos dispostos fora das linhas azuis (desvios entre a amostra original e o resultado obtido) é maior do que a quantidade dos mesmos na forma de onda apresentada para N=256, o que permite concluir que o resultado obtido após a quantização apresenta ligeiros desvios relativamente à amostra inicial. Fato este que se confirmou após a audição do resultado final, onde se consegue perceber um ruído de fundo. Apesar disso continua a ser percetível o conteúdo do ficheiro de áudio.



*Fig. 9 –* Comparação do resultado da quantização com amostra original (N = 16)

Para N=16, foi obtido um Erro Quadrático Médio (MSE) de 0.00134515 e um Peak Signal-to-Noise Ratio (PSNR) de -7.7943. Estes valores confirmam a perceção obtida pela audição do resultado final, uma vez que o MSE é bastante superior à anterior, indicando mais diferenças entre o som original e o quantizado, e o PSNR é muito baixo (negativo) e quanto maior o PSNR, maior a qualidade da reconstrução, pode-se concluir que a qualidade da reconstrução não é a melhor. Conclui-se assim que a amostra obtida por quantização em 256 níveis é a que tem maior qualidade.