

Capítulo 1 – Introdução

Índice

| | |
|---|----|
| Índice..... | 1 |
| 1. Introdução..... | 2 |
| 1.1. Das Ondas Sonoras aos Sinais Eléctricos..... | 2 |
| 1.2. Frequência..... | 4 |
| 1.3. Fase..... | 6 |
| 1.4. Descrição de sinais nos domínios do tempo e da frequência..... | 7 |
| 1.5. Densidade Espectral de Potência..... | 10 |
| 1.6. Transmissão..... | 11 |
| 1.7. Filtros..... | 12 |
| 1.8. Técnicas de Modulação Básicas..... | 17 |
| 1.8.1. Modulação Analógica..... | 17 |
| 1.8.1.1. Modulação de Amplitude..... | 17 |
| 1.8.1.2. QAM – Quadrature Amplitude Modulation..... | 18 |
| 1.8.1.3. Modulação de Frequência..... | 19 |
| 1.8.1.4. Modulação de Fase..... | 20 |
| 1.8.2. Modulação Digital..... | 22 |
| 1.8.2.1. ASK – Amplitude Shift Keying..... | 22 |
| 1.8.2.2. FSK – Frequency Shift Keying..... | 23 |
| 1.8.2.3. PSK – Phase Shift Keying..... | 23 |
| 1.8.3. Modulação por Pulsos Codificados (PCM)..... | 24 |
| 1.8.3.1. Amostragem..... | 24 |
| 1.8.3.2. Quantização..... | 27 |
| 1.8.3.3. Codificação..... | 29 |
| 1.8.4. Modulação por Pulsos Diferencial..... | 32 |
| 1.8.5. Modulação Delta..... | 34 |
| 1.9. Técnicas de Modulação Avançadas..... | 34 |
| 1.9.1. QPSK – QuadriPhase-Shift Keying..... | 34 |
| 1.9.2. M-PSK..... | 35 |
| 1.9.3. QAM revisitado..... | 36 |
| 1.9.4. MSK – Minimum Shift Keying..... | 36 |
| 1.10. Propósitos da Modulação..... | 37 |
| 1.11. Multiplexagem..... | 38 |
| 1.11.1. Multiplexagem por Divisão na Frequência – FDM..... | 39 |
| 1.11.2. Multiplexagem por Divisão no Tempo – TDM..... | 39 |
| 1.11.3. QAM revisitado pela última vez..... | 40 |

1. Introdução

O termo Telecomunicação designa o acto de comunicar à distância. Se, no princípio do século XX a comunicação se cingia à voz, no princípio do século XXI, o volume de informação oriunda de dados e imagem já supera o volume de informação cuja origem é a voz. Assim, os pesos relativos das várias sub-áreas em que se pode dividir a grande área de Telecomunicações têm evoluído enormemente nas últimas décadas para abarcar assuntos que são geralmente designados por multimédia. Actualmente, a diversidade de sistemas que são possíveis de se ligar em rede, e à distância, não pára de aumentar, tornando as Telecomunicações um dos assuntos mais apaixonantes (e atractivos de um ponto de vista de negócio) dos nossos tempos.

De um ponto de vista clássico é usual dividir-se as Telecomunicações em três grandes áreas: Processamento de sinal, Propagação e Redes e Sistemas.

- O Processamento de sinal preocupa-se com a forma como a nossa voz (ou outro tipo de informação) pode ser representada, para depois ser transmitida e recuperada no receptor. Pelo caminho vai sofrer distorções devido aos sistemas que usa, e ao ruído que apanha, e é importante que no receptor a recuperação seja aceitável para permitir a sua compreensão. O Processamento de sinal, não se preocupando em estudar as causas do ruído, estuda as suas consequências e como se pode reconstituir o sinal alterado de um modo tão igual ao original quanto possível.
- A Propagação preocupa-se com o modo como a informação viaja (quer pelo espaço livre, quer por um fio metálico, por uma fibra óptica, etc.), que interferências sofre e como podem ser minimizadas. Inclui o estudo de emissores que tanto podem ser os dispositivos que dão potência ao sinal no início do cabo, como as antenas para transmissão livre.
- Finalmente as Redes e os Sistemas estudam a área de Telecomunicações nos aspectos de ligação de vários componentes para permitir que o sinal que foi gerado, e foi propagado, chegue ao destino pretendido que pode estar num sistema completamente diferente. Trata de todos os equipamentos que façam parte de uma rede no sentido mais lato (comutadores, interligação entre sistemas de mais baixo nível, etc.) e das ligações entre esses equipamentos para permitir um número cada vez maior de facilidades inter-redes para o utilizador¹.

Este texto descreve, de um modo introdutório, a primeira área – o Processamento de Sinal.

1.1. Das Ondas Sonoras aos Sinais Eléctricos

O fenómeno da comunicação pela fala é já bastante conhecido: os músculos das nossas cordas vocais provocam perturbações no fluxo de ar que sai dos pulmões permitindo que moldemos os sons que constituem as palavras. Estas perturbações propagam-se depois no ar, à velocidade do som, e alcançam o destinatário a poucos metros, ou a algumas dezenas de metros de distância.

Um primeiro assunto já de interesse para as Telecomunicações tem a ver com a **potência**. É óbvio que se falarmos calmamente com alguém perto de nós, falamos num tom baixo. Mas se quisermos contactar alguém a uns cem metros de distância teremos de gritar, ou de produzir uma perturbação com maior potência. Claramente, existe um limite para essa potência e não nos é possível criar uma perturbação sonora tal que sejamos ouvidos a, por exemplo, dez quilómetros de distância. Simplesmente, as perturbações produzidas vão perdendo a sua força com a distância – a **potência** vai decrescendo em função da distância.

Outro fenómeno bem conhecido dos alunos do primeiro ano de uma Universidade é a **relação sinal-ruído** (ou mantendo-nos na nomenclatura desta introdução, a relação perturbação-ruído). Principalmente nas aulas teóricas, se a conversa entre os alunos das primeiras filas do anfiteatro for animada, os alunos

¹ Por exemplo, há uns anos atrás um grande objectivo poderia ser pagar a conta da electricidade através de um telemóvel. E este ano?

das últimas filas não ouvem o que o Professor diz. O sinal produzido pelo Professor sofre interferências pelo caminho chegando ininteligível ao final do anfiteatro. Existem dois modos de se solucionar este problema, como é sabido: ou o Professor produz um sinal com mais potência, falando mais alto, de tal modo que a potência do sinal é tão maior do que a do ruído que o torna ainda perceptível nas últimas filas; ou os alunos das primeiras filas param de falar (de produzir ruído, portanto) possibilitando que o sinal que é produzido pelo Professor, com a mesma potência que anteriormente, chegue às últimas filas pois, a relação da sua potência com a potência do ruído aumentou, por se ter diminuído a potência do ruído. Resumindo, tanto num caso como no outro a relação entre a potência do sinal e a potência do ruído aumentou. Em conclusão, para além da potência do sinal, outro aspecto importante na comunicação é também como ela se relaciona com a potência do ruído.

Deixando a problemática importantíssima das aulas teóricas, retornemos à impossibilidade de falar com outra pessoa distanciada de dez quilómetros. Para que tal seja possível, é preciso fazer uso de meios artificiais, pois as nossas cordas vocais têm um limite prático. Um modo de o executar é passar a perturbação sonora para uma perturbação eléctrica, transmiti-la e depois recuperá-la. Outro modo, que será abordado mais adiante, seria inventar uns altifalantes tão potentes que criassem um sinal sonoro de grande potência para poder ser ouvido a dez quilómetros. Será que funcionaria? Repare que, este segundo modo acaba por ser semelhante ao primeiro, pois tem a passagem para eléctrico, amplificação e a passagem para sonoro, diferindo apenas no modo como o sinal se propaga (ou electricamente até aos mini-altifalantes a dez quilómetros, ou por ondas sonoras). Antes de prosseguir para novos conceitos, exploremos um pouco mais os conceitos relacionados com estas passagens de e para eléctrico, amplificações e transmissões.

Primeiro, vamo-nos concentrar na passagem de sonoro para eléctrico e na recuperação final para sonoro. Os equipamentos que as executam designam-se genericamente de **transdutores**. Basicamente, na passagem para eléctrico existe uma membrana que é afectada pela perturbação sonora e se move para trás e para a frente produzindo indução eléctrica, e consequentemente, corrente eléctrica. Está-se a falar, evidentemente, de um microfone. No caso da recuperação basta que exista um circuito electromagnético que, pela passagem de corrente, produza afastamentos e aproximações de uma membrana. Estes movimentos, por sua vez, produzem as perturbações sonoras equivalentes. Está-se agora a falar dos altifalantes. Como pode ver as coisas são muito parecidas. Será que um altifalante serve de microfone? Tente ligar uns “*headphones*” na entrada de som de um computador, a fazer de microfone e veja o que acontece...

O conhecimento da forma destas perturbações ao longo do tempo, tanto eléctricas como sonoras, vai permitir escolher as melhores “ferramentas” matemáticas para estudarmos a área das Telecomunicações e evoluir nela. Qual será o melhor método para as tratar? Vamos ver primeiro como elas são...

Se imaginarmos que um altifalante (ou uma membrana do microfone) está na sua posição de repouso num certo ponto, as perturbações provocam que a membrana se desloque para trás e para a frente desse ponto. Se monitorássemos todas essas posições ao longo do tempo, poderíamos traçá-las num gráfico em função do tempo (admitindo, por exemplo, que afastamentos para um lado do ponto de repouso são negativos, e para o outro são positivos). A Figura 1 representa uma possível hipótese.

A primeira observação da Figura 1 é que, se quiséssemos representar esta linha por uma função do tempo, teríamos uma tarefa muito complicada, pois ela não se assemelha a nenhuma função matemática simples conhecida (como seno, logaritmo, exponencial, etc.). A seu tempo vamo-nos preocupar com o assunto de como representar funções destas...

Admitindo que a relação entre os sinais eléctricos e os movimentos das membranas é de proporcionalidade directa, podemos aceitar que afinal a linha da Figura 1 também representa o sinal eléctrico correspondente à perturbação sonora do emissor, aparte um factor de escala. Podemos encará-la como o valor da tensão, ou da corrente. Então, a nossa tarefa está terminada, pois basta enviar esse sinal eléctrico por uma linha (ou pelo ar), o receptor apanha-o e envia-o para o altifalante. Infelizmente, a transmissão provoca sempre alterações no sinal, tornando-o, por vezes irrecoverável. Os motivos

principais são o ruído, as interferências e as características do meio de transmissão (espaço, metal, ou fibra óptica). Estas alterações estão muito dependentes da **frequência** do sinal, que vai ser o tema abordado a seguir. Antes, e para deixar de vez as ondas sonoras, é sabido da experiência do dia a dia que a solução de potentes altifalantes para falar a dez quilómetros é ineficiente pois: (a) existe muito ruído e interferências do “mesmo tipo”: o vento, o barulho das folhas das árvores, dos automóveis, etc.; e (b) existe a própria atenuação das perturbações com a distância (a atenuação do meio) que impedem a comunicação.

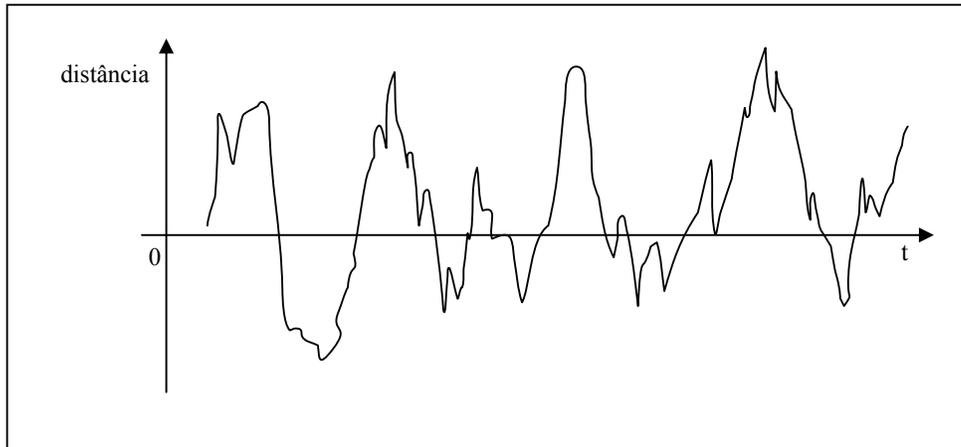


Figura 1

Representação da posição da membrana de um altifalante ao longo do tempo, em relação ao seu ponto de repouso.

Os meios de transmissão eléctricos (se estivermos a falar de um fio metálico) ou electromagnéticos (se estivermos a falar do espaço) têm também fenómenos do “mesmo tipo” que vão afectar a progressão do sinal:

- (1) Vai existir ruído eléctrico e electromagnético. Por exemplo, perto de uma central de produção de energia, ou de um posto de transformação eléctrico. O efeito é semelhante à conversa dos alunos das primeiras filas das aulas teóricas, e mais uma vez, é importante saber qual a relação que existe entre a potência do sinal e a do ruído.
- (2) Vai existir atenuação. Do mesmo modo que as perturbações sonoras perdem potência com a distância, também a potência do sinal eléctrico vai sendo cada vez menor com a distância. O mais grave é que o sinal é afectado de maneiras diferentes nas várias frequências que possui. Isto é, certas frequências perdem mais potência com a distância do que outras.

1.2. Frequência

Vamos então abordar o conceito de **frequência** de um sinal. Frequência, de um modo geral, serve para quantificar quantas vezes se repete um fenómeno periódico por unidade de tempo. À unidade de “ciclos por segundo” chama-se Hertz (Hz). Por exemplo, um motor de automóvel a andar num ritmo não esforçado pode estar trabalhar a 3600 rotações por minuto. Este automóvel estaria, assim, a trabalhar a 60 Hz.

Outro exemplo, mais na área de Telecomunicações, tem inspiração na música. Uma nota musical pura (se é que isso existe) tem apenas uma frequência. Designar-se-ia por um sinal de tom único, ou monocromático. O cepticismo sobre a existência de notas puras é propositado pois (felizmente) os instrumentos musicais não conseguem produzir uma frequência única pura. Existem sempre outras frequências (harmónicas) com muito menos potência do que a principal. Ao conjunto destas frequências chamamos **timbre** e são elas que nos permitem distinguir um Dó de um piano do Dó de um violino. O único instrumento feito pelo homem que consegue produzir quase uma só frequência é o diapasão, que normalmente é feito para produzir um Lá. Mesmo assim, o diapasão também não é ideal por duas razões:

produz ainda outras frequências, com uma potência muito mais baixa, e a frequência não se mantém sempre constante pois o Lá vai-se extinguindo com o tempo.

A membrana de um altifalante que reproduzisse fielmente o som de um diapasão ideal (que também não existe, portanto) estaria a fazer um movimento periódico sinusoidal com a frequência dessa nota, como está mostrado na Figura 2 (o fenómeno de extinção da nota com o tempo traduzir-se-ia numa atenuação dos valores máximos da sinusóide ao longo do tempo, como está mostrado na linha a ponteados na figura).

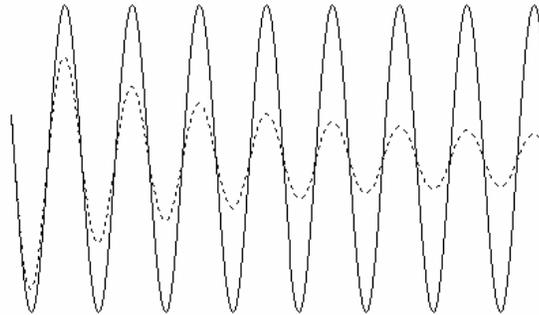


Figura 2

Representação da posição da membrana de um altifalante ao longo do tempo, em relação ao seu ponto de repouso, ao reproduzir o som de um diapasão ideal.

O gráfico da Figura 2 representa uma função coseno, pois, por convenção, usa-se o coseno em Telecomunicações, em vez do seno. Uma expressão matemática possível para descrever o valor do sinal, o valor da tensão do sinal, $v(t)$, seria

$$v(t) = A \cos(Bt)$$

em que A representa a amplitude máxima do sinal e B é um valor tal que reflecte a frequência com que os movimentos da membrana se repetem. A expressão anterior é totalmente válida para descrever a figura, mas o B não viria em ciclos por segundo (Hz), mas sim em *radianos por segundo* (rad/s), a **frequência angular**. A letra mais usada para a frequência angular é o ω , pelo que a expressão anterior é normalmente escrita do seguinte modo

$$v(t) = A \cos(\omega t)$$

Voltando ao exemplo do automóvel, é estranho falarmos de Hertz para descrever as rotações do motor. Preferimos falar em rpm (rotações por minuto). Em Telecomunicações, por sua vez, é mais usual utilizar-se o Hertz para descrever a frequência dos sinais, do que a frequência angular em radianos por segundo. Mais adiante vai-se ver que não é apenas por tradição, mas simplifica algumas coisas. A relação entre a frequência angular e a frequência é $\omega = 2\pi f$. Assim, a expressão anterior toma agora a forma

$$v(t) = A \cos(2\pi f_0 t)$$

Como o coseno tem um período de 2π , é claro na expressão anterior que em cada segundo existem f_0 períodos completos do coseno, portanto, f_0 ciclos por segundo, ou f_0 Hz.

Se construíssemos um diapasão para vibrar numa nota musical mais grave, com uma frequência menor, teríamos um comportamento como o mostrado na Figura 3. Se outro diapasão vibrasse numa nota mais aguda, frequência maior, seria qualquer coisa como o mostrado na Figura 4.

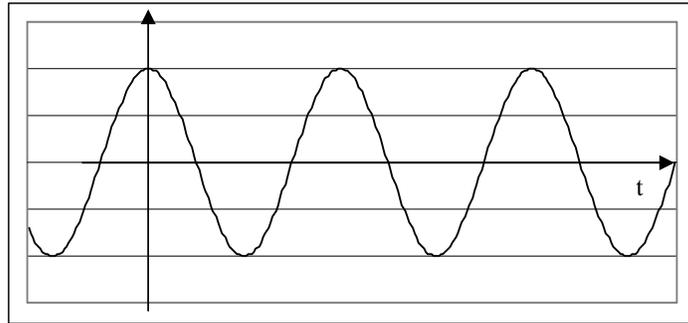


Figura 3
Representação da posição da membrana de um altifalante ao longo do tempo, em relação ao seu ponto de repouso, reproduzindo uma nota musical mais grave do que a da Figura 2.

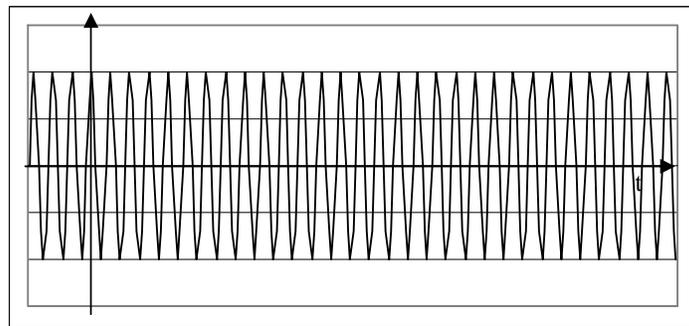


Figura 4
Representação da posição da membrana de um altifalante ao longo do tempo, em relação ao seu ponto de repouso, reproduzindo uma nota musical mais aguda do que a da Figura 2.

1.3. Fase

Pegando na representação da Figura 3, se deslocarmos a onda um pouco para a direita, isso significaria que a nota musical começou a ser tocada um pouco mais tarde. A Figura 5 tem uma representação dos dois sinais. É o mesmo sinal, com a mesma frequência, apenas atrasado um pouco no tempo.

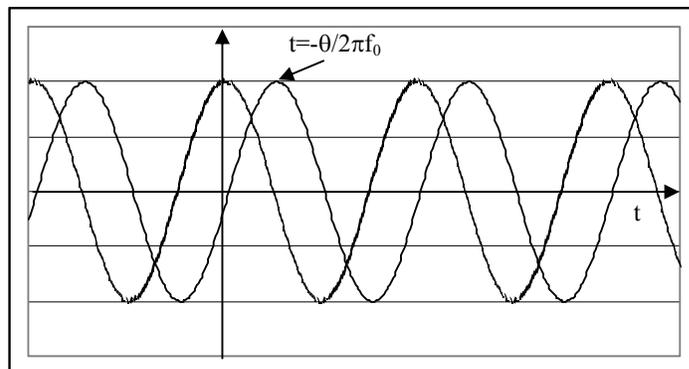


Figura 5
Representação da mesma nota musical começada a tocar numa certa altura e um pouco mais tarde.

Como distinguir um sinal do outro? Será que um sinal foi simplesmente deslocado do pequeno intervalo que é óbvio quando se olha para a figura, ou foi deslocado de um período mais aquele bocado? Ou de dois períodos mais aquele bocado? Mais uma vez, por convenção, usa-se apenas o bocado mínimo entre as duas ondas, e a expressão da onda deslocada para a direita é a seguinte

$$v(t) = A \cos(2\pi f_0 t + \theta)$$

O termo novo que aparece, θ , designa-se por fase do sinal. **A fase representa o facto de o pico não estar alinhado com a origem dos tempos, e ocorrer em $t = -\theta / 2\pi f_0$.** Aliás a expressão anterior é geral, e representa também a onda original, para a qual $\theta = 0$.

1.4. Descrição de sinais nos domínios do tempo e da frequência

Quando se analisam a Figura 2 e a Figura 5, os sinais nelas representados variam a sua amplitude ao longo do tempo. No entanto, cada um deles tem sempre a mesma frequência ao longo do tempo, uma amplitude máxima e a indeterminação de saber se começam mais para a esquerda ou para a direita é decidida pelo valor da fase. Por outras palavras, estes sinais estariam completamente descritos se disséssemos

a frequência,
a amplitude máxima e
a fase

A isto chama-se a **descrição do sinal no domínio da frequência**, ou **espectro do sinal**. É uma descrição em função da variável frequência. Às descrições dos sinais nos moldes da Figura 2 e da Figura 5 chamam-se **descrições dos sinais no domínio do tempo**, pois mostram o valor da amplitude do sinal ao longo do tempo.

Por exemplo, a descrição gráfica do sinal da Figura 3 no domínio da frequência é representada por **dois gráficos**: um para a **frequência** e outro para a **fase**, como mostra a Figura 6. O primeiro gráfico mostra que o sinal tem apenas uma frequência, a f_0 , e a amplitude é A (a amplitude máxima do coseno da Figura 3). O segundo gráfico mostra que na frequência f_0 , (a única que o sinal tem) a fase vale θ . Ao primeiro gráfico chama-se a **representação de amplitude** (ou **espectro de amplitude**) e ao segundo a **representação de fase** (ou **espectro de fase**)².

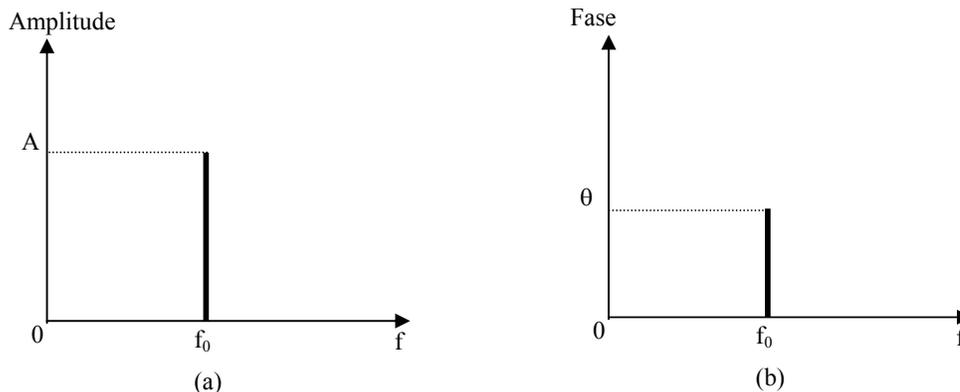


Figura 6
Representação do espectro do sinal (a) representação da amplitude; (b) representação da fase.

² Normalmente, a representação da amplitude é sempre positiva, pois os possíveis valores negativos são absorvidos pela fase usando a expressão

$$-A \cos(2\pi f_0 t + \theta) = A \cos(2\pi f_0 t + \theta \pm 180^\circ)$$

Na Figura 7 estão representadas duas notas musicais produzidas pelos tais diapasões ideais, uma mais grave e outra mais aguda. Para simplificar, ambas têm a mesma amplitude, e a mesma fase, diferente de zero. Já agora, e como curiosidade, o ouvido humano não é sensível à fase. Isto é, se a nota mais grave for começada a tocar primeiro do que a mais aguda, nós não conseguimos distinguir o som conjunto das duas notas, do som conjunto produzido quando se toca a nota mais aguda primeiro.

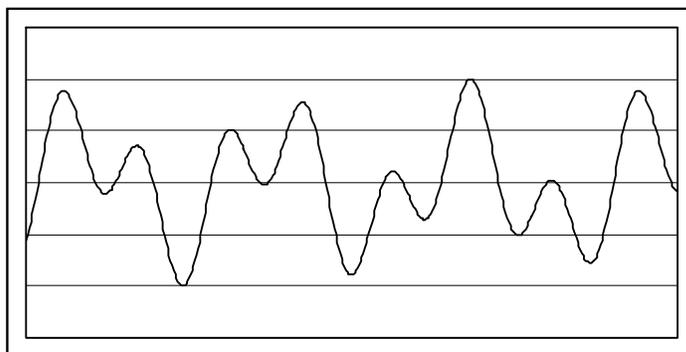


Figura 7
Representação de duas ondas com frequências f_0 e f_1 .

A função da Figura 7 parece complicada. No entanto são apenas duas frequências puras. Se as representássemos pelo seu espectro seria apenas como está mostrado na Figura 8. Como se pode ver, mesmo que fosse apenas usado para as representações de funções, o espectro pode ser uma alternativa muito interessante. Mas a importância do espectro excede largamente a representação de funções, como se irá ver.

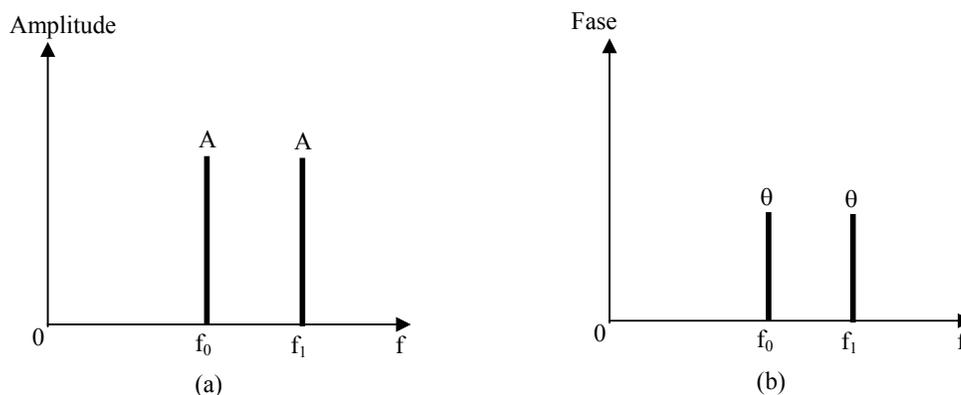


Figura 8
Espectro das duas ondas cuja representação no tempo está mostrada na Figura 7

Na continuação desta Introdução vai-se deixar para segundo plano o espectro de fase. Não é completamente relevante para a explicação dos conceitos base e assim ficaremos mais focalizados nos aspectos de amplitude.

Como curiosidade preste atenção à Figura 9. Na parte (a) está mostrado um sinal com seis frequências em que quatro delas têm uma amplitude muito pequena comparadas com as duas frequências que lhe estão próximas. Na parte (b) essas quatro frequências foram retiradas. Será que o ouvido humano distingue o sinal (a) do sinal (b)? A resposta é não! Quando temos uma frequência com uma amplitude muito pequena ao pé de outra com uma amplitude muito grande, não conseguimos perceber a frequência com amplitude pequena. Então poderemos ver-nos livre dessas frequências e conseguir guardar um sinal num ficheiro bem mais pequeno. É precisamente isso que faz o MP3! Espero que nesta altura, o leitor já esteja completamente rendido à utilidade de pensarmos os sinais em frequência. Muitos fenómenos ficam triviais quando se pensa no domínio da frequência. Outros exemplos irão aparecer mais

adiante. Já agora, a decisão de deitar fora uma frequência tem a ver com quão pequena é a amplitude ao lado de outra grande e quão perto é o “ao lado”. Estes limiares variam com um parâmetro de qualidade do MP3. Por isso existem as mesmas músicas com ficheiros de tamanhos diferentes...

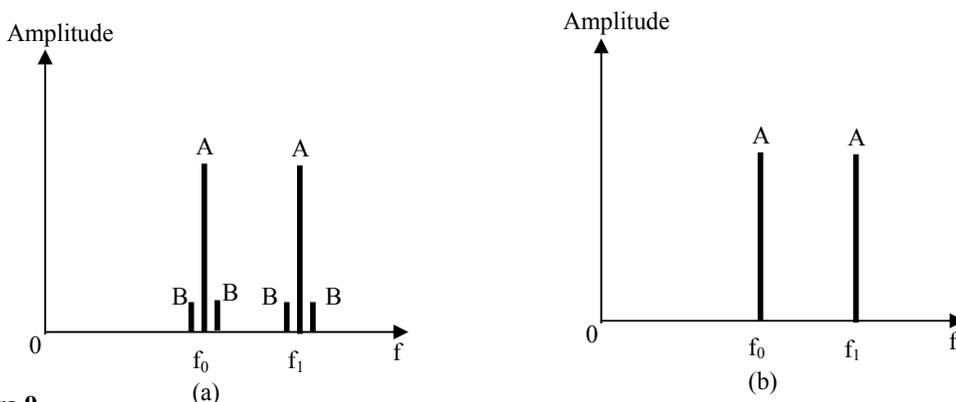


Figura 9
Um sinal com seis frequências (a) e outro sinal com apenas as duas frequências com maior amplitude (b)

Voltando à Figura 8, ela já nos permite falar de outro conceito. O que a sua representação de amplitude mostra é que na frequência f_0 existe uma certa amplitude, na frequência f_1 existe outra amplitude, e não existem mais nenhuma componente de frequência no sinal. Por exemplo, o sinal não tem componentes numa frequência f_2 (se imaginarmos que f_2 está a meio de f_0 e f_1). A nível de frequência não existem outras componentes, ou seja, em função da frequência o sinal não está definido para outros valores de frequência que não sejam f_0 e f_1 . Este tipo de sinal diz-se que tem um **espectro discreto** (não contínuo) e isso acontece sempre que o **sinal no tempo é periódico e infinito** (como os cosenos)³.

A fala humana, evidentemente, não é periódica nem infinita. O seu espectro não é discreto, mas sim **contínuo**. A Figura 10 mostra os valores do espectro de amplitude da fala humana. Como se vê, a função está definida continuamente entre os 400 e os 4.000 Hz, para todos os valores da frequência. Como curiosidade, uma cantora de ópera seria capaz de produzir sons com uma frequência maior do que a maioria das outras pessoas. Já agora, uma pergunta para responder com experiências na aula prática fica já aqui: Até que frequência é que o ouvido humano consegue ouvir? É bom saber isto para não sermos levados pelos fabricantes de aparelhagens de alta fidelidade... A Figura 10 mostra também que o sinal emitido quando **falamos** é limitado na frequência a pouco mais de 4.000 Hz para um ser humano comum. Não existe nada depois disso!... Nos sistemas de telecomunicações que vamos construir não vale a pena considerar frequências maiores do que essa. Podiam-se ter telefones de alta fidelidade, mas não é economicamente viável.

Aliás como curiosidade o rádio AM, como o conhecemos, vai até aos 5.000 Hz, o que faz com que a voz seja de qualidade, mas a música seja de fraca qualidade. Já as estações FM vão até aos 15.000 Hz permitindo uma difusão de música de boa qualidade.

Uma ideia bem ousada seria pensar na função representada na Figura 10 como o limite de um grande número de frequências f_0, f_1, \dots, f_n representativas de frequências puras. A função contínua da Figura 10 seria o limite de um espectro discreto do tipo do da Figura 8!...Muito mais frequências e cada vez mais juntinhas... Se fôssemos por este caminho, então a fala humana poderia ser **representada no tempo** como o somatório (ou, no limite, um integral) de um grande número de cosenos, cada um com a amplitude dada pelo valor da função no espectro de amplitude da Figura 10 para o seu valor de frequência. Se tivéssemos também o espectro de fase, então saberíamos o atraso que cada coseno teria relativamente aos outros. Será que há vantagem em pensar deste modo? A resposta virá ao longo da disciplina...

³ Claro que não existem sinais infinitos na realidade. Eles têm de começar algures no tempo. No entanto, se a sua frequência for tão grande comparada com o período de observação pode-se considerar que eles são, de facto, periódicos e infinitos

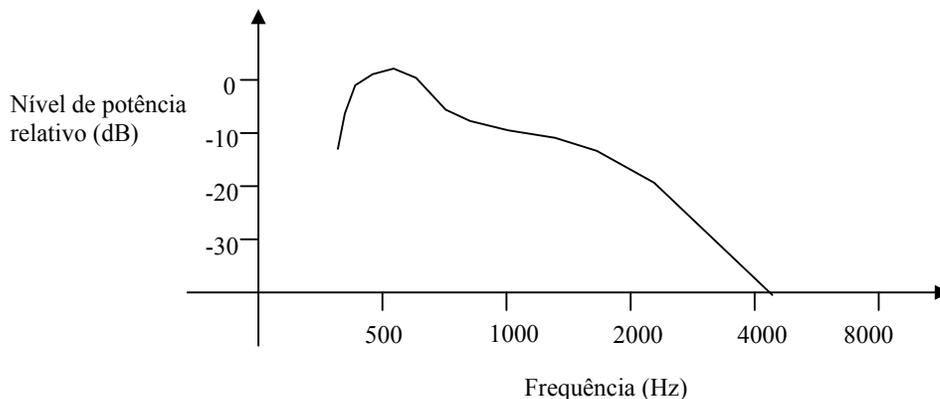


Figura 10
Representação do espectro de amplitude da fala humana

1.5. Densidade Espectral de Potência

Um aspecto um pouco mais técnico, mas de grande utilidade para as Telecomunicações, é saber como a potência do sinal se distribui para cada frequência que o sinal tem. A simplicidade do seu cálculo é grande, uma vez tendo a descrição do sinal na frequência dada pelo espectro. Como se sabe, a potência dissipada numa resistência R em função da tensão é dada por⁴

$$p(t) = \frac{v^2(t)}{R}$$

Isto é, a potência é proporcional ao quadrado da amplitude do sinal. Se considerarmos ainda que se tem uma resistência de 1 ohm (que é comum considerar-se em Telecomunicações) a potência tem a mesma expressão quer se use a tensão ou a corrente para descrever o sinal. O que é importante nesta altura da descrição é que a relação é quadrática. Ora, o que se verifica é que a distribuição da potência do sinal pela frequência é o quadrado do valor do espectro de amplitude. A essa função dá-se o nome de **densidade espectral de potência**. Para o caso da fala humana, a potência dos sons que produzimos quando falamos tem uma função que é o quadrado da função mostrada na Figura 10. Vê-se que existe muito mais potência nas frequências até 1.000 Hz e depois valores menores até 4.000 Hz.

Mais uma vez como curiosidade, poderíamos ser levados a cortar as frequências mais altas da nossa fala, pois até têm pouca potência... Assim, conseguíamos usar equipamentos que transportassem apenas até 2.000 Hz e poupava-se bastante. No entanto, existe outro factor importante que é a inteligibilidade do sinal. Na nossa fala as frequências menores são usadas quando dizemos vogais e as maiores quando produzimos as consoantes. Mesmo com pouca potência é importante elas lá estarem, senão não compreendemos o discurso. Um exemplo sobre a linguagem escrita torna este problema mais visível. Na escrita a importância das consoantes também é fundamental quando comparada com as vogais. Por exemplo, a palavra *espectro* sem vogais seria *spectr* o que faz lembrar vagamente a palavra real. Se usássemos apenas as vogais seria quase impossível saber que palavra seria. Assim, mesmo com pouca potência, é importante considerar a fala humana até pouco mais de 3.000 Hz (normalmente até 3.400 Hz).

A densidade espectral de potência da fala humana é descrita pelo quadrado dos valores da função mostrada na Figura 10. Esta “ligação” entre o espectro de amplitude e a densidade espectral de potência

⁴ Do mesmo modo se poderia considerar a potência em função da corrente – $p(t)=Ri^2(t)$. Tal como com a tensão, a potência é proporcional ao quadrado da amplitude do sinal.

faz com que o espectro de amplitude seja sempre uma referência mais usada do que o espectro de fase (que, no entanto, também é bastante importante como se irá ver ao longo da disciplina).

1.6. Transmissão

Como se disse atrás, os vários meios de transmissão não “conduzem” todas as frequências de igual modo. Certas frequências são pouco atenuadas por quilómetros, enquanto outras frequências sofrem atenuações muito grandes logo em centenas de metros. Se tivermos um cabo com uns dez quilómetros (tipicamente a distância de um telefone doméstico a uma central telefónica em zonas não urbanas) podemos pensar que as componentes de frequência que são muito atenuadas não chegam, pura e simplesmente, à central. Vejamos um exemplo.

Pensando apenas no espectro de amplitude, a Figura 11 mostra na parte de cima, seis notas musicais no tempo e na frequência. Na parte de baixo está ilustrado o que seria o sinal no final do cabo se as duas maiores frequências não existissem (tivessem sido atenuadas pelo meio de transmissão). Claramente, não temos o mesmo sinal e o que ouviríamos não se pareceria muito com o que foi produzido.

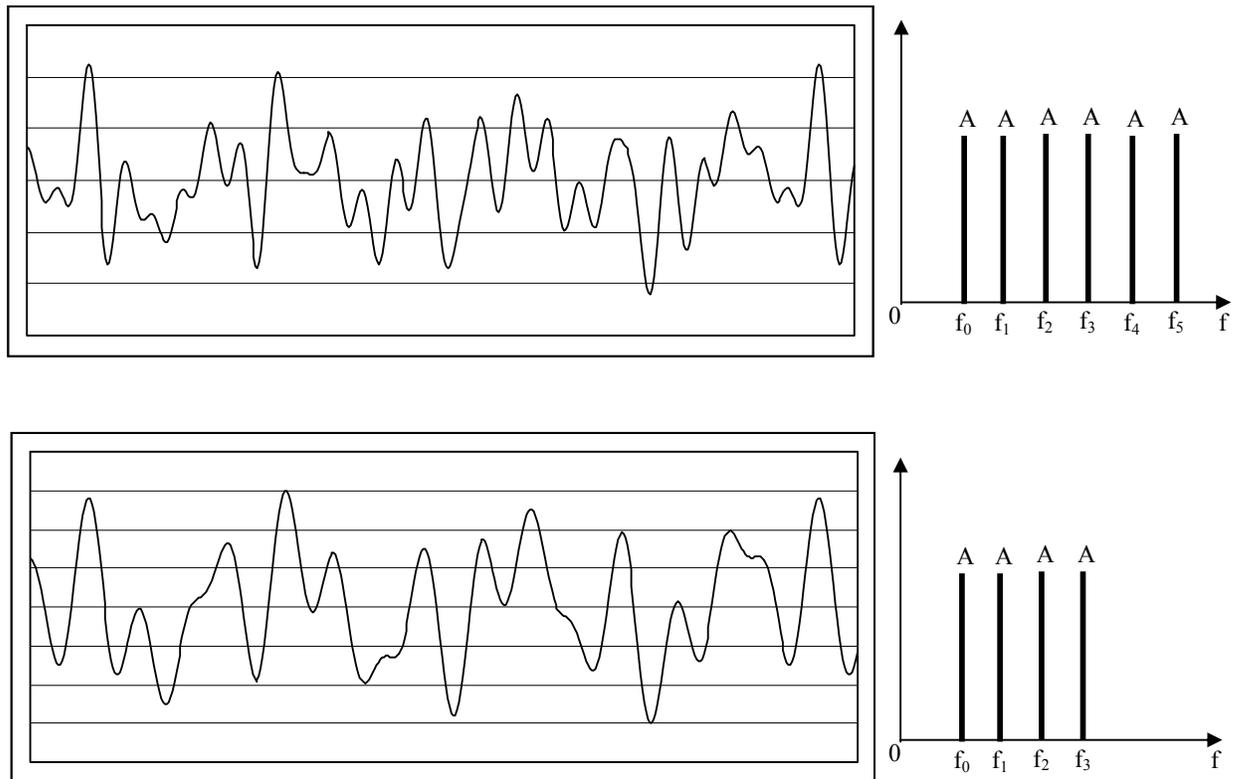


Figura 11

Representação do sinal no tempo e do espectro de amplitude com seis frequências e com apenas as quatro frequências mais baixas.

A consequência é que se a distância for muito grande, pode acontecer que já não se consiga recuperar o sinal no destino. Mesmo na gama das frequências onde a atenuação é pequena, ela existe. Se a distância for muito grande a potência do sinal pode ficar tão fraca que se assemelha à potência do ruído. Para evitar isto, no caso das distâncias a percorrer serem muito grandes, têm de se usar amplificadores de

sinal de tantos em tantos quilómetros. A distância exacta entre amplificadores tem a ver com a atenuação do meio e a potência do ruído. Cá nos aparece, mais uma vez, a relação sinal-ruído...

No caso concreto do espaço livre, de que são exemplos a difusão de rádio e de televisão, as frequências de interesse começam nas chamadas frequências médias, MF (*Medium Frequency*), para as emissões de rádio em onda média; passam pelas muito altas frequências, VHF (*Very High Frequency*), para as emissões de rádio em FM e televisão; e vão até às ultra altas frequências, UHF (*Ultra High Frequency*) para a televisão e os telemóveis. Em termos de frequências começa-se em 500 kHz, passa-se por 100 MHz e vai-se até 1 ou 2 GHz.

Acima do UHF, as utilizações são já de feixes de microondas e de satélite. Abaixo dos 500 kHz não existe “condução” (ou propagação) suficientemente eficiente. Este cenário cria um problema: “*Como transmitir então o sinal eléctrico da nossa voz que começa nos 400 Hz e vai até aos 4.000 Hz?*”. Se transmitirmos o sinal tal como ele é à saída do microfone, isto é, se ligássemos o microfone à antena, ele degradar-se-ia poucos metros depois de deixar a antena que o tentasse transmitir...

A solução será transladar, de algum modo, o sinal para as frequências de interesse (as tais MF, VHF ou UHF), transmiti-lo nessas frequências e passar de volta para as frequências de 400 Hz a 4.000 Hz no receptor. O processo de passar o sinal para outra frequência tem o nome genérico de **modulação**. Atenção, não confundir **modulação** com *modelação*, como é habitual em 90% dos alunos. Não se está a tentar achar um modelo para o que quer que seja, mas sim fazer uma translação de frequência.

1.7. Filtros

Antes de se entrar na explicação do processo de modulação, foquemos um pouco mais os diferentes efeitos dos meios de transmissão em função da frequência e abordemos também os dispositivos chamados **filtros**. Os filtros são dispositivos que actuam selectivamente na frequência. Isto é, pensando no domínio da frequência, os filtros modificam o sinal que lhes entra, num sinal à saída que não tem amplitudes em certas frequências. O sinal à saída é igual ao sinal à entrada para certas frequências, enquanto que para outras é nulo.

Um filtro que cortasse todas as altas frequências a partir de um certo valor teria um comportamento idêntico ao do canal representado na Figura 11. Ou, dito de outro modo, aquele canal de transmissão acaba por ter um comportamento equivalente ao de um filtro. Como se pode ver das partes direitas dessa figura, a descrição do efeito do filtro na frequência é trivial (cortou as frequências f_4 e f_5).

Uma descrição no tempo do efeito do filtro é bem mais complicada. Basta tentar ver como se pode chegar na parte superior esquerda da figura para a parte inferior esquerda. O estudo no tempo baseia-se na resposta do filtro a um impulso especial que é o mínimo impulso temporal que se pode ter, chamado de **impulso unitário** e mostrado na Figura 12. Este impulso, denominado $\delta(t)$, tem um valor diferente de zero em $t=0$ e um valor de zero para todos os outros valores de t . A sua representação é tão estreita, que normalmente se desenha como uma seta no ponto em que ele é diferente de zero, como está representado na parte direita da Figura 12.

Se introduzirmos um impulso unitário à entrada do filtro a excitação do filtro fica confinada ao instante $t=0$. Qualquer saída observada depois de $t=0$ é característica desse filtro em particular, pois o sinal de entrada já se extinguiu, e é chamada de **resposta impulsiva**. É como se nós impuséssemos uma rajada abrupta de energia em $t=0$ e observássemos o filtro a estabilizar sozinho. Por esta razão a resposta impulsiva é também denominada de **resposta natural**, designando-se por $h(t)$. A Figura 13 mostra uma possível resposta impulsiva de um filtro F1.

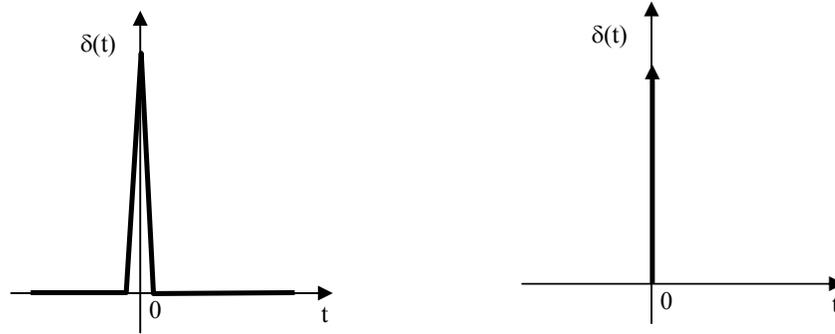


Figura 12
Representação no tempo de um impulso unitário.

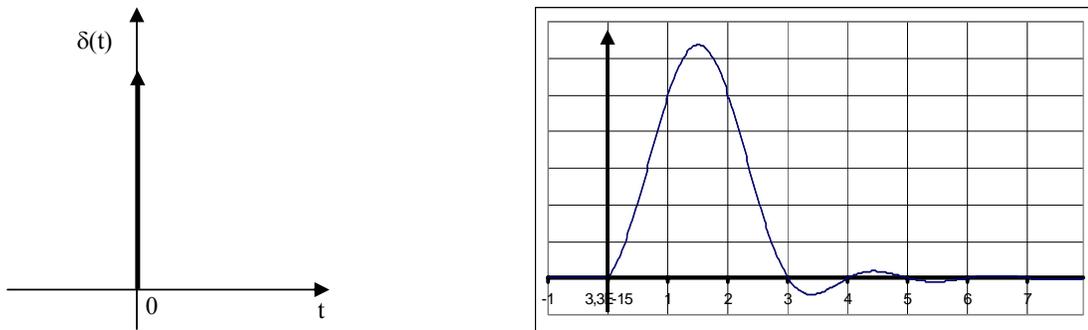
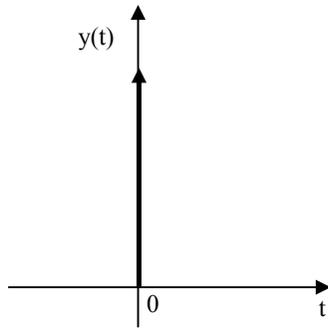


Figura 13
Resposta impulsiva de um filtro F1.

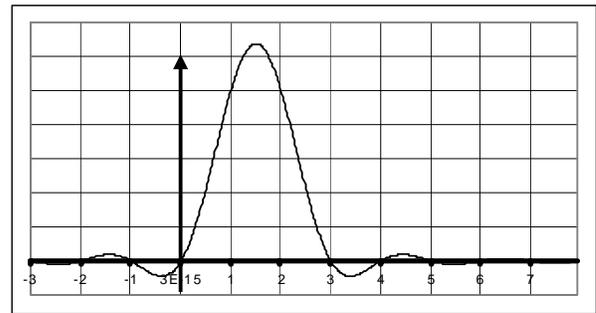
As respostas impulsivas podem ser muito diferentes dependendo dos circuitos. A Figura 14 mostra quatro tipos de respostas impulsivas de interesse. O caso (a) mostra um circuito sem memória (bastante monótono!); o caso (b) mostra um filtro não causal pois começa a responder ainda antes da entrada estar activa (um filtro destes não tem realização prática, como é óbvio); o caso (c) mostra um filtro instável porque $h(t)$ cresce ilimitadamente; finalmente o caso (d) mostra um filtro causal, estável e com memória.

Porque são importantes o impulso unitário e a resposta impulsiva para se descrever o efeito de um filtro no tempo?

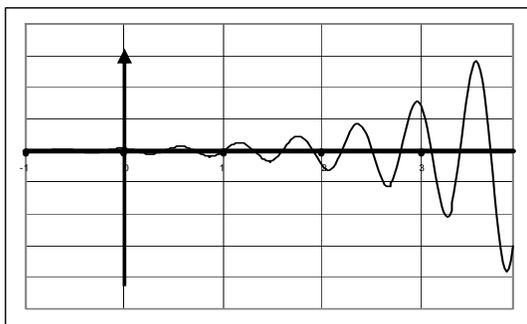
Se o filtro permitir que a resposta à soma de duas excitações seja igual à soma das respostas de cada excitação individualmente, então a resposta, por exemplo, a dois impulsos unitários espaçados no tempo seria a curva soma das duas curvas individuais da Figura 15.



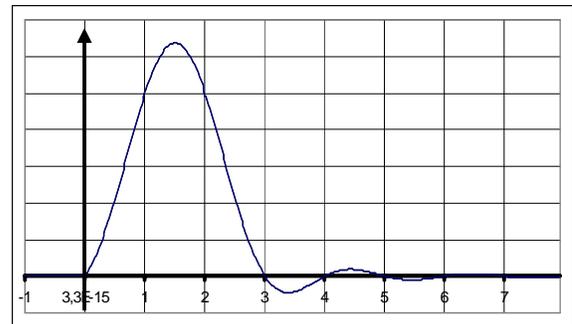
caso (a)



caso (b)



caso (c)



caso (d)

Figura 14
Quatro tipos de respostas impulsivas

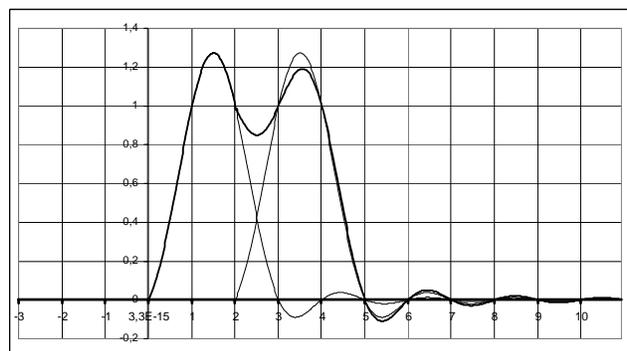
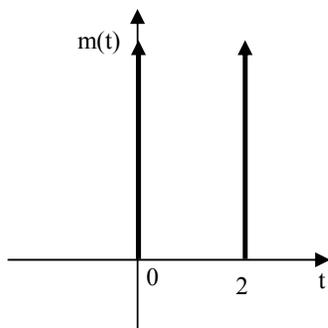


Figura 15
Resposta a dois impulsos unitários espaçados no tempo

Percebe-se na Figura 15 que existem três tempos envolvidos: o tempo de excitação, o tempo de resposta e o tempo de memória do sistema (a parte em que ele ainda está a reagir a algo do passado). Como incorporar então a memória do sistema e generalizar para qualquer sinal de entrada?

A resposta a um impulso unitário em $t=0$ é

$$h(t)$$

A resposta a um impulso unitário em $t=\tau$ é

$$h(t - \tau)$$

Se o impulso tiver uma magnitude de a_0 e a_τ então as respostas vêm

$$a_0 h(t) \quad e \quad a_\tau h(t - \tau)$$

respectivamente.

O passo que resta para se ter a resposta a um sinal qualquer à entrada é agora trivial.

Imaginemos que à entrada se tinha um sinal qualquer, como o representado na Figura 16. Este sinal pode ser visto como a soma (aliás integral) de uma sequência de impulsos unitários tão “juntinhos” cuja magnitude de cada um é descrita pelo valor de $x(t)$ para cada t . Então a resposta é a soma (aliás integral) de todas as respostas impulsivas diferenciadas no tempo pesadas pelas respectivas magnitudes. Isto é,

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau$$

A esta expressão dá-se o nome de **integral de convolução**. Como se vê, o estudo no domínio do tempo é bem mais complicado do que no domínio da frequência...Aqui fica mais um argumento da importância e simplicidade de se considerarem as descrições no domínio da frequência quando se querem estudar certos fenómenos.

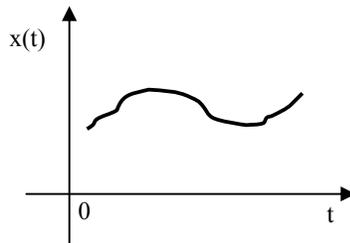


Figura 16
Sinal $x(t)$ de entrada

Do mesmo modo que falamos de descrições no domínio da frequência para descrever as características de um sinal, também podemos descrever as características dos filtros na frequência por meio do espectro. Aliás, para além dos filtros podemos descrever as características de qualquer circuito na frequência por meio do espectro. O nome genérico que se usa para qualquer circuito, filtros incluídos, é o termo **sistema**. Assim, o espectro de amplitude do filtro de que a Figura 11 poderia ser um exemplo está representado na Figura 17. Vê-se que existe uma região não nula na frequência, que significa que os componentes daquelas frequências são afectados com um ganho de C , e uma região nula em que os componentes dessas frequências são, pura e simplesmente, anulados. O valor C pode ser unitário significando que o filtro não introduz nenhum ganho nas frequências que deixa passar. Se C for diferente de um, o filtro introduz uma amplificação ($C>1$) ou uma atenuação no sinal à saída ($C<1$). A um filtro deste género, que deixa passar todas as frequências a começar em zero até um certo valor, a partir do qual, elimina todas as frequências, designa-se por **filtro passa-baixo**. À frequência de transição chama-se frequência de corte, f_c .

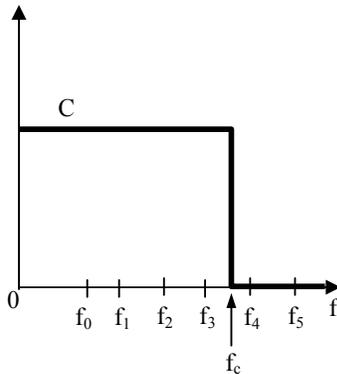


Figura 17
Espectro de amplitude de um filtro passa-baixo.

Podemos imaginar outros tipos de filtros:

- (a) uns que eliminem todas as frequências desde zero até um certo valor, e deixem passar todas as frequências a partir daí. A este tipo de filtros chama-se **filtros passa-alto**;
- (b) outros que só deixem passar as frequências num certo intervalo, que se designam por **filtros passa-banda**;
- (c) ou ainda outros que eliminem todas as frequências de um certo intervalo, que se designam por **filtros rejeita-banda**.

Espectros de amplitude destes três tipos de filtros estão representados na Figura 18.

Nos circuitos de telecomunicações, o papel dos filtros vai ser muito importante para eliminar partes indesejáveis dos sinais que são produzidas como efeitos secundários dos vários processos pelos quais os sinais passam.

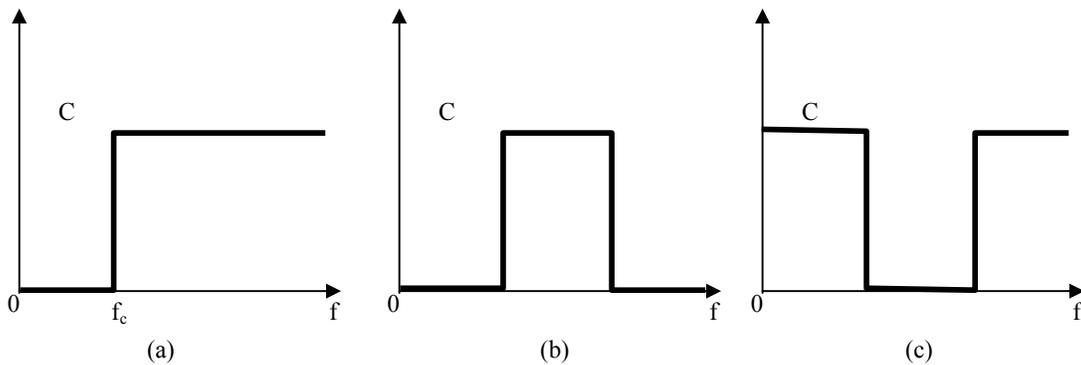


Figura 18
Espectros de amplitude de (a) um filtro passa-alto, (b) um filtro passa-banda, e (c) um filtro rejeita-banda.

1.8. Técnicas de Modulação Básicas

Depois desta incursão pelos filtros, voltemos ao tema da modulação. Existem, basicamente, três tipos de modulação: a modulação analógica, a modulação digital (que de um ponto de vista de conceitos pode ser considerada como um caso particular da modulação analógica) e a modulação por pulsos. Depois existem algumas variantes que combinam estas básicas, ou fazem as coisas de um modo mais específico.

1.8.1. Modulação Analógica

As modulação analógica, por sua vez, tem três variantes. Elas assentam num princípio comum: o uso de uma frequência pura, com um certo valor, que se designa por **onda portadora**. A onda portadora é uma sinusóide e, como tal, não tem informação muito relevante – é como os dias da semana. Se disséssemos que tínhamos nascido numa Quarta-feira, não estávamos a dizer muita informação sobre a nossa idade...

É necessário, então, alterar a monotonia da sinusóide portadora com a informação do nosso sinal (**sinal modulante**). Podemos fazer isso, como se disse, de três maneiras:

- alterando a amplitude da portadora – modulação de amplitude (AM);
- alterando a frequência da portadora – modulação de frequência (FM);
- ou alterando a fase da portadora – modulação de fase (PM).

No conjunto de variantes mais evoluídas das modulações básicas, existem modulações que alteram mais do que um destes parâmetros ao mesmo tempo. O objectivo é conseguir mais eficiência (e complicação). A onda resultante dessa alteração chama-se **sinal modulado**, ou **onda modulada**.

Convém não esquecer que no destino tem de se inverter o processo, fazer a **desmodulação**, para depois enviar o sinal na frequência certa para o altifalante para podermos ouvir. Se enviássemos um sinal de 800 kHz para um altifalante, por exemplo, mesmo que mecanicamente fosse possível, os nossos ouvidos não conseguiriam captar a mensagem. Atente-se, por exemplo, aos assobios para cães...

No caso de alterarmos a amplitude da onda portadora, o espectro do sinal modulado fica igualzinho ao espectro do sinal inicial, mas deslocado para a frequência da portadora. Por este facto, designa-se este processo de modulação como **modulação linear**. Nos casos da alteração da frequência, ou da fase, os espectros das ondas moduladas são diferentes do espectro do sinal original (ocupam mais frequências) e designam-se estes processos como **modulação exponencial**. Como contrapartida de um maior uso de frequências, obtém-se nestes últimos casos uma maior imunidade ao ruído.

O processo de modulação oferece também um modo de se **multiplexarem** vários sinais no mesmo meio. Este fenómeno é conhecido de todos nós na possibilidade de sintonizarmos os nossos rádios em várias estações, ou as nossas televisões em vários canais. As várias estações de rádio e de televisão estão todas a transmitir ao mesmo tempo. Elas usam simplesmente ondas portadoras de diferentes frequências no mesmo meio de comunicação, o espaço, acordadas com uma entidade reguladora – a ANACOM, no caso de Portugal. A “separação” entre elas é de tal modo que não têm conflitos umas com as outras.

1.8.1.1. Modulação de Amplitude

Começemos, então, pela alteração da amplitude. A ideia é aumentar a amplitude da onda portadora quando a tensão do nosso sinal é grande, e diminuí-la quando a tensão do nosso sinal é pequena. No receptor temos de ter um circuito que reaja de modo recíproco – gere uma onda que tenha a sua amplitude em função da amplitude da onda portadora no momento.

A Figura 19 mostra uma onda modulante, $m(t)$, e a onda modulada, $s(t)$, que tem uma frequência da portadora de f_c (c de *carrier*). Para simplificar esta explicação, a onda modulante é sempre não negativa (depois hão-de estudar-se os outros casos). Já agora, outro aspecto que os alunos se devem habituar é que o sinal modulante (que poderia ser a nossa voz, ou uma música) é normalmente também uma sinusóide.

Haveremos de ver a razão de se utilizarem também sinusóides para os sinais modulantes. O que interessa aqui referir é alertar os alunos para esta situação. Pode ser confuso falar de sinais modulantes e portadoras sendo todas sinusóides...

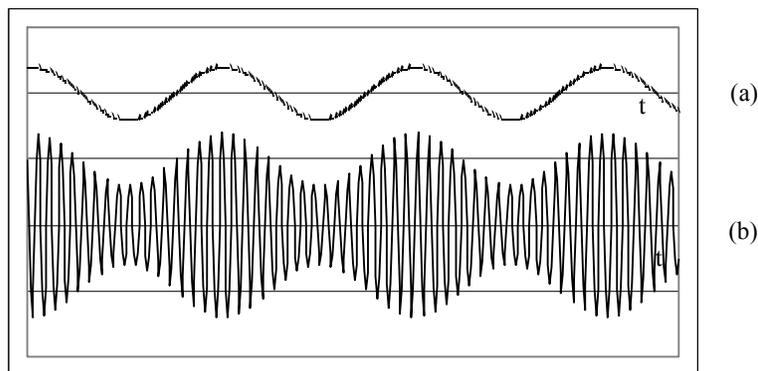


Figura 19
Formas das ondas (a) modulante, e (b) modulada para a modulação de amplitude

A Figura 20 mostra os espectros das ondas modulante e modulada. Como se vê, a forma do espectro (que foi inventada para a figura) fica inalterada, passando simplesmente para a frequência da portadora⁵

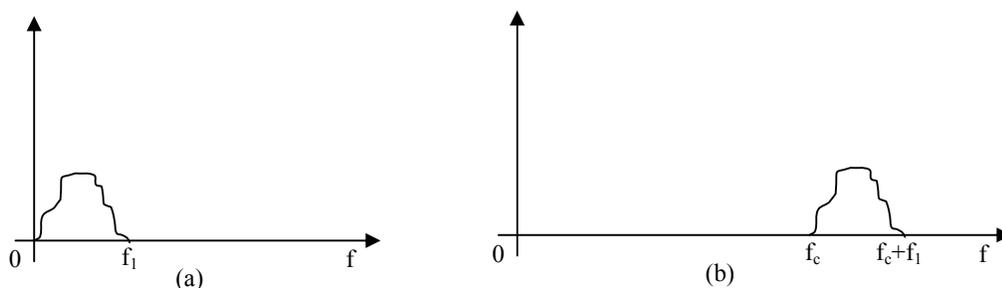


Figura 20
Espectros de amplitude da (a) onda modulante, (b) onda modulada para a modulação de amplitude.

O problema da transmissão fica, assim, resolvido, pois a onda modulada tem uma frequência desde f_c a f_c+f_1 , e consegue ser propagada no espaço. No receptor há que voltar a recolocá-la nas frequências desde 0 a f_1 e passá-la a sonoro (se estivermos a falar de voz). Tem-se, portanto, de deslocar o espectro outra vez para o local da parte (a) da Figura 20.

1.8.1.2. QAM – Quadrature Amplitude Modulation

Uma pequena variante da modulação de amplitude é o QAM (*Quadrature Amplitude Modulation*). A ideia é usar uma portadora coseno para transmitir um sinal modulante $m_1(t)$ e uma portadora seno para transmitir outro sinal modulante $m_2(t)$. O nome de quadratura advém do facto do seno e do coseno serem ortogonais. À primeira vista isto parece impossível, pois as ondas vão misturar-se as duas...

A Figura 21 tenta mostrar as duas ondas numa representação vectorial para se visualizar melhor o facto de elas estarem mesmo separadas. No entanto, na prática, isso não existe e as ondas vão como está

⁵ Na continuação da disciplina vamos ver que as coisas não são bem assim, e aparece outra forma do lado esquerdo de f_c . Mas, por agora, cinjamo-nos a esta explicação mais trivial.

mostrado na Figura 22. Uma confusão... A pergunta que fica nesta introdução é como é que o receptor consegue a partir de uma onda como a da Figura 22 retirar os sinais $m_1(t)$ e $m_2(t)$, tão nitidamente distintos na Figura 21. Havemos de conseguir responder a isto no final da disciplina...

Em termos de representação na frequência, o espectro é muito parecido com o da Figura 20, em que a frequência máxima, f_s , é a frequência maior dos dois sinais $m_1(t)$ e $m_2(t)$.

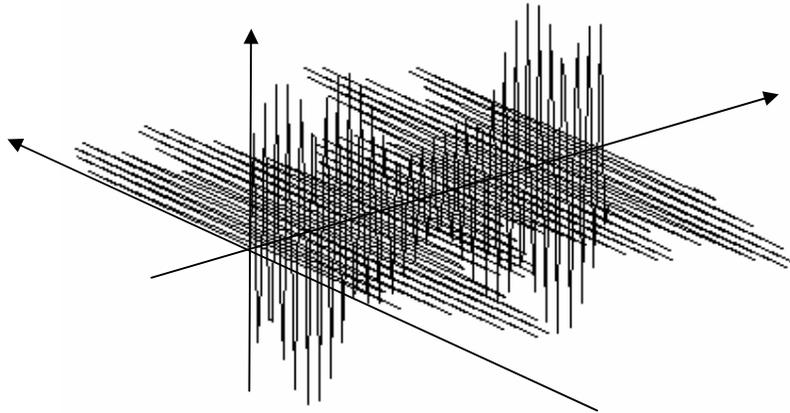


Figura 21

Uma onda QAM com dois sinais, mostrada vectorialmente entre seno e coseno

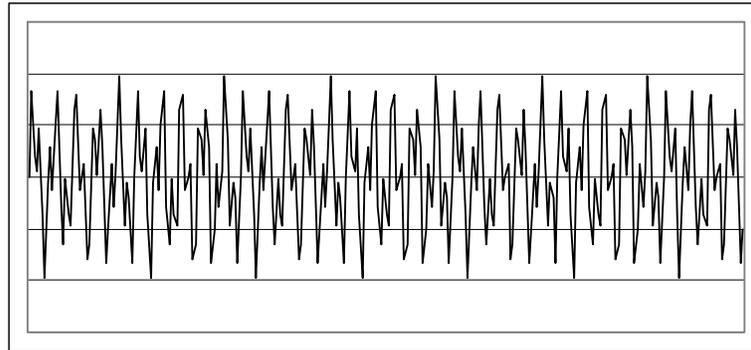


Figura 22

Uma onda QAM com dois sinais mostrada no tempo

1.8.1.3. Modulação de Frequência

A ideia base da modulação de frequência é usar a portadora numa frequência f_c , e alterar-lhe a frequência em função da onda modulante. Se a amplitude da onda modulante é grande, a frequência da portadora deve subir bastante. Se a amplitude da onda modulante é pequena, a frequência da portadora deve ser só ligeiramente superior a f_c . No caso da amplitude da modulante ser negativa, o processo é semelhante, mas para frequências inferiores a f_c .

A Figura 23 mostra novamente uma onda modulante sinusóidal e a onda modulada respectiva. Atenção que as diferenças de frequência da onda modulada são muito ligeiras e não são perceptíveis ao olho. A figura está extremamente exagerada.

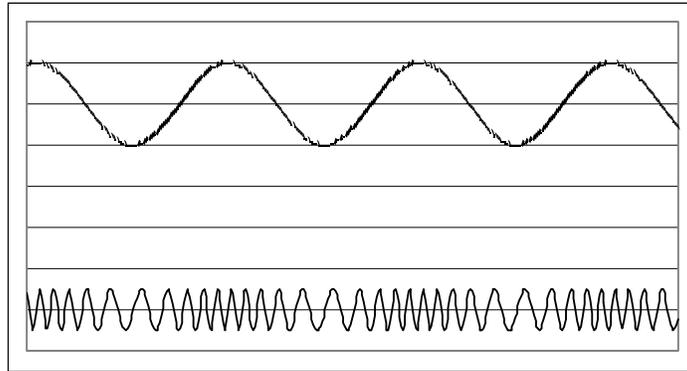


Figura 23
Formas das ondas (a) modulante, e (b) modulada para a modulação de frequência

No receptor, para se voltar a obter a onda original, tem de se usar um dispositivo que produza uma onda com uma amplitude proporcional ao afastamento, em cada momento, que a frequência da onda modulada tem relativamente a f_c . De um ponto de vista espectral, já não existe uma relação linear, como se disse. A Figura 24 mostra uma hipotética relação espectral entre as ondas modulante e modulada.

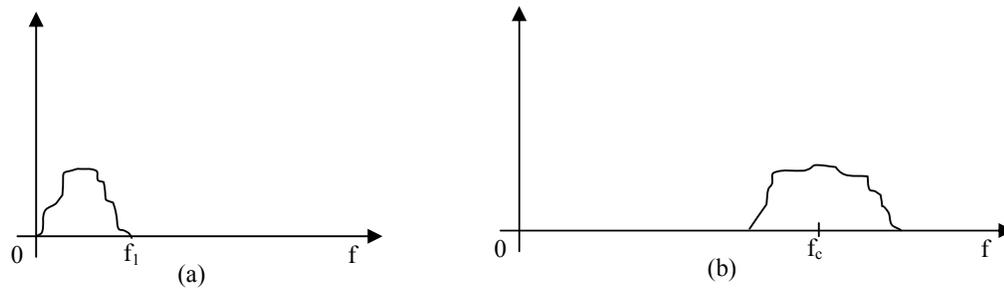


Figura 24
Espectros de amplitude da (a) onda modulante, (b) onda modulada para a modulação de frequência.

1.8.1.4. Modulação de Fase

A ideia base da modulação de fase é usar a portadora numa frequência f_c , que mantém a sua frequência inalterada, e muda-se a fase em função da onda modulante. Na prática é como se atrasássemos a onda mais ou menos consoante a amplitude da onda modulante fosse maior ou menor. No caso de ela ser negativa adiantava-se a onda portadora mais ou menos. Tal como para a modulação de frequência, a onda modulada tem sempre a mesma amplitude máxima. Isto é bastante bom pois a sua potência é sempre constante e não varia com a onda modulante, como na modulação de amplitude.

Existe uma relação muito estreita entre a modulação de frequência e a modulação de fase. Ela baseia-se na relação entre a frequência instantânea (a frequência que a onda tem num momento muito curto) e a fase. A relação é a de derivada. Mais propriamente, sendo $f_i(t)$ a frequência instantânea e $\theta(t)$ a fase

$$f_i(t) = \frac{1}{2\pi} \frac{d\theta(t)}{dt}$$

Esta relação é importante pois uma modulação de fase é simplesmente uma modulação de frequência da derivada da onda modulante, $m(t)$ ⁶.

No receptor vai ser preciso um circuito que a cada momento calcule o atraso ou o adiantamento da onda modulada relativamente à fase zero e produza uma onda com amplitude proporcional a essa diferença. Evidentemente, pode-se usar simplesmente um receptor de modulação de frequência, obter a onda e depois integrá-la para ter a verdadeira onda modulante.

Os exemplos que se estão a dar de formas de onda usam frequências puras. Ora a derivada de um coseno é um seno (portanto, um coseno com uma diferença de fase) o que significa que a forma de onda da onda modulada em fase vai ser muito parecida com a forma de onda da onda modulada em frequência mostrada na Figura 23. De facto assim é. A Figura 25 mostra novamente uma onda modulante de tom único e a onda modulada respectiva. Mais uma vez foi extremamente exagerada. Notar que agora as frequências maiores não estão alinhadas com o máximo da onda, mas nas partes onde a “subida” é acentuada (estão com o máximo da onda derivada que seria o seno. Isto é, se derivássemos a onda modulante bastar-nos-ia fazer uma modulação de frequência).

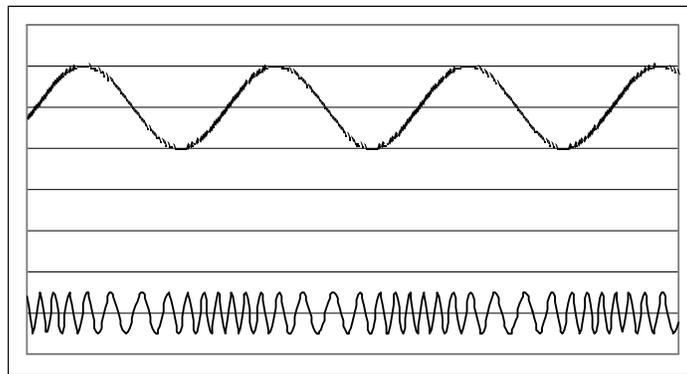


Figura 25
Formas das ondas (a) modulante, e (b) modulada para a modulação de fase

De um ponto de vista espectral, também já não existe uma relação linear, como se disse, entre o espectro da onda original e o espectro da onda modulada. Embora os espectros da modulação de frequência e da modulação de fase sejam diferentes, não vamos ligar muito a isso nesta parte introdutória da disciplina, e a Figura 26 serve perfeitamente para ilustrar uma possível relação entre os espectros.

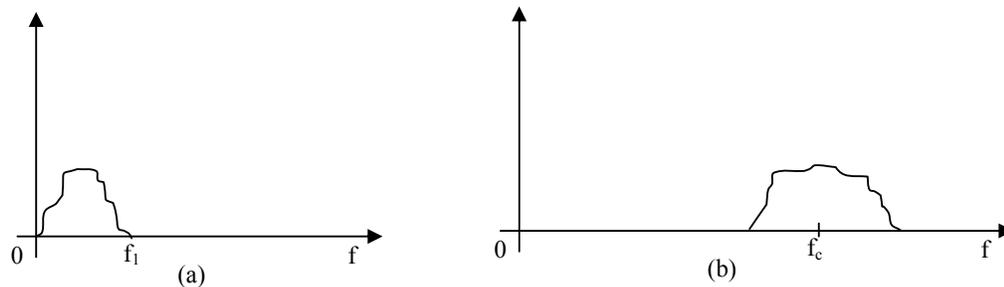


Figura 26
Espectros de amplitude da (a) onda modulante, (b) onda modulada para a modulação de fase.

⁶ Podíamos ter explicado as duas modulações de maneira inversa e a modulação de frequência seria simplesmente uma modulação de fase do integral da onda modulante.

1.8.2. Modulação Digital

Cada vez mais se usa informação em formato digital (o que significa formato digital será explicado mais adiante). As razões desta escolha serão estudadas durante a disciplina. Por agora, interessa pensar que o nosso sinal vai ter apenas dois valores no tempo, em vez de poder tomar qualquer valor. Assim, as modulações anteriores poderiam ser usadas tal como foram pensadas para analógico em que agora haveria apenas as seguintes diferenças:

- Na modulação de amplitude a amplitude máxima da onda modulada poderá ter apenas dois valores – A_1 ou A_2 .
- Na modulação de frequência, a onda modulada poderá ter apenas uma frequência f_1 ou uma frequência f_2
- Na modulação de fase, a onda modulada poderá não ter qualquer atraso ou ter um atraso de 180° .

Só para não simplificar demasiado o problema, poderíamos pensar que em vez de digitalizarmos o nosso sinal em binário (em que só são possíveis dois valores – 0 e 1) poderíamos tê-lo digitalizado em quaternário, com quatro valores de tensão (em que cada valor corresponderia a dois bits), ou digitalizado com oito valores (em que cada valor corresponderia a três bits), etc. Então as modulações teriam mais do que dois valores possíveis cada uma. Mas isso são apenas extensões ao que já foi dito.

Outro pensamento é se decidirmos mudar mais do que um parâmetro ao mesmo tempo (amplitude, frequência e fase) como se verá mais adiante. Nestas técnicas de modulação mais avançadas faz-se um uso intensivo do conhecimento do número de valores que se está a usar para otimizar todo o processo. Neste ponto vamos-nos cingir apenas a um sinal digital binário (dois valores, um bit) e mudança de apenas um parâmetro.

Como na modulação digital só se pode alternar entre dois valores (ou entre um número fixo de valores) decidiu-se chamar às várias modulações o nome de *Shift Keying*. Assim, os três tipos de modulações analógicas já estudadas têm o nome em digital de, respectivamente, *Amplitude Shift Keying* (ASK), *Frequency Shift Keying* (FSK) e *Phase Shift Keying* (PSK).

Um ponto **muito importante** na comunicação digital que é muitas vezes esquecido é o tempo. Na comunicação analógica as coisas vão acontecendo... Na comunicação digital é muito importante saber quanto tempo dura cada símbolo, neste caso um bit. Assim, é como se a transmissão ficasse igual durante um símbolo (bit), para depois mudar e continuar igual no próximo símbolo (bit), e assim por diante.

1.8.2.1. ASK – Amplitude Shift Keying

Na modulação digital de amplitude o símbolo binário 1 é representado pela transmissão da onda portadora sinusoidal de amplitude A e frequência f_c , durante o tempo que se decidiu que o símbolo deve durar. Por exemplo, se estivermos a transmitir a 100 Kbit por segundo cada símbolo dura 10μ segundos. O símbolo binário 0 é representado por não se transmitir sinal nenhum durante o seu tempo. Uma sequência binária 0 1 0 1, etc. está representada na Figura 27. O receptor tem apenas que detectar se num dado intervalo de tempo correspondente a um símbolo se está a transmitir a portadora, ou não, para decidir se é um 1 ou um 0.

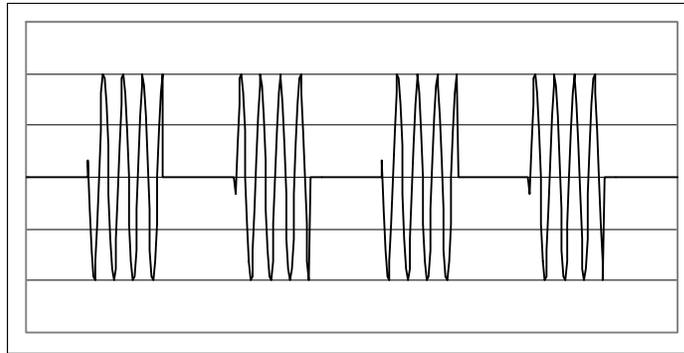


Figura 27

Forma da onda modulada em ASK para a sequência binária 0 1 0 1 0 1 0 etc.

1.8.2.2. FSK – Frequency Shift Keying

Num sistema FSK usam-se duas portadoras sinusóidais de frequências f_1 e f_2 . O símbolo binário 1 é representado pela transmissão de uma das frequências e o símbolo binário 0 é representado pela transmissão da outra frequência, durante o tempo de cada símbolo. A forma de onda da mesma sequência binária anterior está representada na Figura 28. O receptor pode ser construído de diversas maneiras. Uma maneira possível é o uso de dois filtros muito estreitos centrados em f_1 e f_2 . Consoante haja sinal na saída de um dos filtros ou do outro, assim se transmitiu o símbolo binário 1 ou 0.

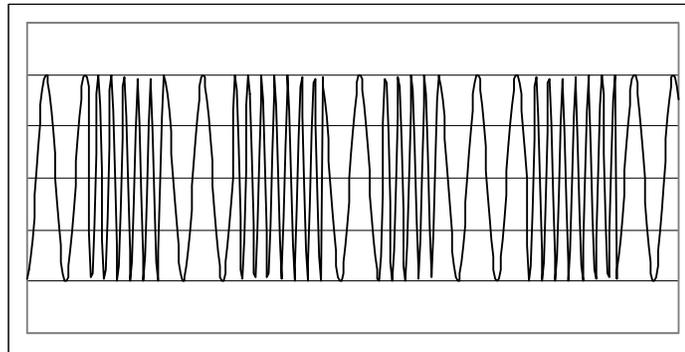


Figura 28

Forma da onda modulada em FSK para a sequência binária 0 1 0 1 0 1 0 etc.

1.8.2.3. PSK – Phase Shift Keying

Num sistema PSK a mesma onda portadora de amplitude A e frequência f_c é usada para representar ambos os símbolos binários 0 e 1. A diferença é que o símbolo binário 0 (por exemplo) usa a portadora com uma fase de -180° , enquanto que no símbolo binário 1 se transmite a portadora com fase 0, durante o tempo do símbolo, tal como está ilustrado na Figura 29.

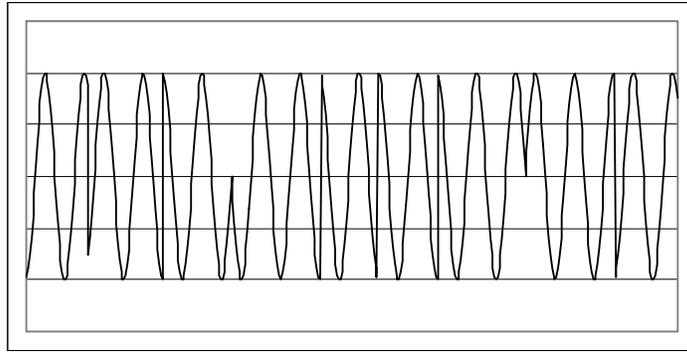


Figura 29

Forma da onda modulada em PSK para a sequência binária 0 1 0 1 0 1 0 etc.

Nota-se na figura os saltos de fase cada vez que muda o símbolo binário. Os saltos poderiam ser mesmo discretos e a onda “aparecer” de repente com outra amplitude. Na modulação FSK a fase é contínua e a onda não tem este tipo de saltos (Figura 28). O receptor de PSK é mais complicado pois tem de detectar se a portadora que está a receber tem o atraso, ou não⁷.

Como se pode ver, nos três tipos de modulação digital, as frequências utilizadas podem ter valores tais que estejam na gama das frequências que podem ser propagadas no meio de transmissão escolhido. À medida que se aumenta o ritmo binário, o tempo do símbolo diminui tornando a tarefa do receptor decidir o que está a receber mais crítica e exigindo circuitos cada vez mais precisos.

1.8.3. Modulação por Pulsos Codificados (PCM)

A modulação por pulsos é iminentemente diferente dos tipos de modulação anteriores. Vamos começar por analisar brevemente a modulação por pulsos codificados – *Pulse Code Modulation* (PCM). Outras duas variantes serão brevemente descritas a seguir. O PCM consiste em três operações: *amostragem*, *quantização* e *codificação*.

1.8.3.1. Amostragem

A modulação PCM é, também, digital. Um assunto que ainda não foi focado é o que é um **sinal digital** e como se digitaliza um sinal analógico. O processo que se vai descrever baseia-se no **teorema da amostragem**. O primeiro passo é converter um **sinal contínuo** no tempo para um **sinal discreto** no tempo.

O teorema da amostragem diz que é possível passar um sinal contínuo para um sinal discreto e conseguir recuperar novamente o sinal contínuo a partir do discreto desde que se obtenham amostras a um ritmo superior a um certo valor.

Se tivermos um sinal que tenha frequências até W Hz, tem de se amostrar esse sinal a, pelo menos, um ritmo de $2W$ amostras por segundo. Em situações ideais até nem se necessita de ter amostras a ritmos superiores, pois não trazem mais informação. Se se amostrar a um ritmo inferior a $2W$ amostras por segundo, por exemplo a Z amostras por segundo (em que $Z < W$), perdem-se as componentes de frequência do sinal entre Z e W .

⁷ De facto, é fácil deduzir o esquema do receptor de PSK. No entanto, necessita de um tratamento matemático que só será estudado durante a disciplina.

Assim, de tempos a tempos, mais concretamente de $1/2W$ segundos em $1/2W$ segundos, amostra-se o sinal contínuo, como está ilustrado na Figura 30, obtendo-se uma sequência de valores $\{m(n/2W)\}$ que tem toda a informação contida em $m(t)$.

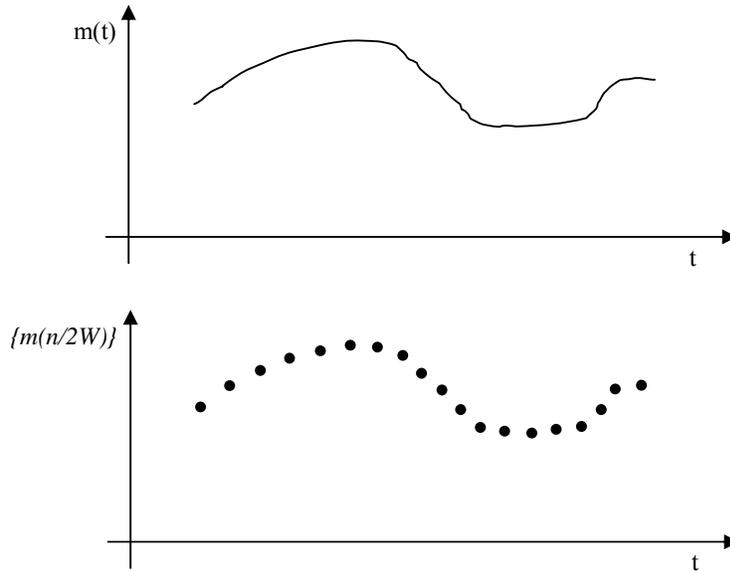


Figura 30
Amostragem de um sinal $m(t)$

No receptor vai-se ter de fazer a operação inversa – passar de uma sequência discreta para um sinal contínuo. Se se cumprir o ritmo de amostragem conveniente no emissor, esta operação inversa não provoca nenhuma alteração ao sinal original e obtém-se o mesmo sinal.

É interessante analisar um pouco mais esta operação de amostragem para se ver os problemas que podem existir na prática. Da maneira como foi explicado, as amostras são instantâneas. Isto é, o valor da função é obtido num único ponto no tempo. Em termos teóricos isso é possível utilizando uma sequência de impulsos unitários (ver Figura 12) em que cada impulso só é diferente de zero no ponto de amostragem. Em termos de frequência, o espectro de um impulso unitário é constante e igual a 1, como está ilustrado na Figura 31.

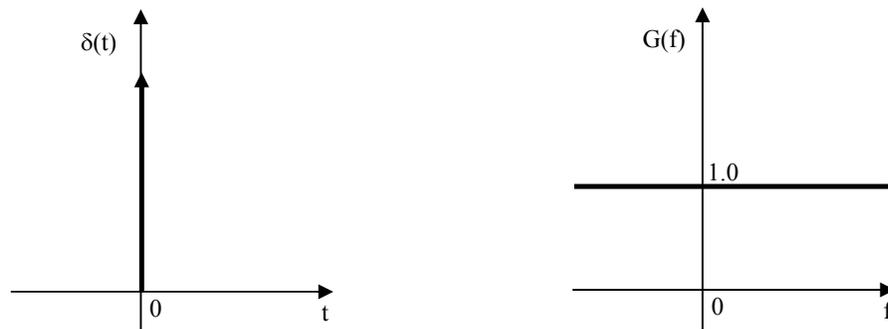


Figura 31
Forma de onda no tempo de um impulso unitário, e espectro de amplitude (na frequência) do impulso unitário.

Imagine-se agora um sinal com uma frequência máxima de W Hz, e que foi amostrado a uma frequência de amostragem, f_s , igual a $2W$ Hz. A Figura 32 mostra o espectro do sinal original (a sua forma foi inventada só para este exemplo), assim como o espectro que resulta do processo de amostragem. A razão de ser dos espectros terem uma parte negativa que é simétrica da parte positiva ficará mais clara na

continuação da disciplina. Como se vê, depois da amostragem, a forma repete-se indefinidamente na frequência. Na prática, só se usa a primeira forma (a cinzento) pois ela contém toda a informação. As outras formas têm a mesma informação, mas usam frequências mais altas. Consegue-se usar só a primeira forma filtrando o sinal amostrado com um filtro passa-baixo, e transmitindo para o receptor o que se obtém à saída do filtro. O espectro da primeira forma é semelhante ao espectro original, tendo apenas um factor de escala de diferença (de magnitude). Esta semelhança é mais uma prova de que o processo de amostragem e de recuperação não introduz nenhuma alteração ao sinal.

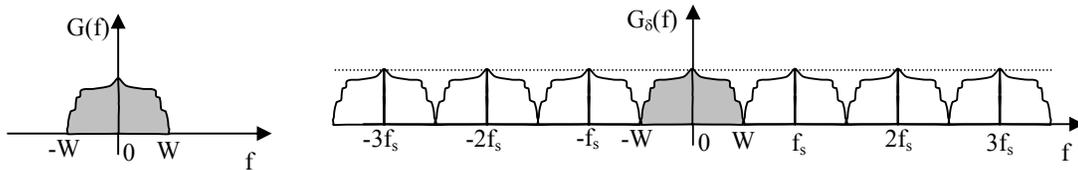


Figura 32

Espectro do sinal original contínuo limitado na frequência a W Hz, e espectro da sua versão amostrada com impulsos unitários a um ritmo de amostragem igual a $2W$ Hz.

Na realidade os impulsos unitários não existem! Não existe um sinal que seja diferente de zero apenas num ponto do tempo. A amostragem é efectuada com pulsos muito curtos, como o da Figura 33. O espectro de amplitude de um pulso já não é uma constante na frequência, mas uma função da forma da parte direita da Figura 33. Quanto mais estreito for a duração do pulso, mais “largo” é o lóbulo principal da parte direita da Figura 33. Isto é natural, pois quanto mais estreito for o pulso, mais se parece com o impulso unitário, e mais o lóbulo principal na parte direita da figura se parece com a recta da parte direita da Figura 31.

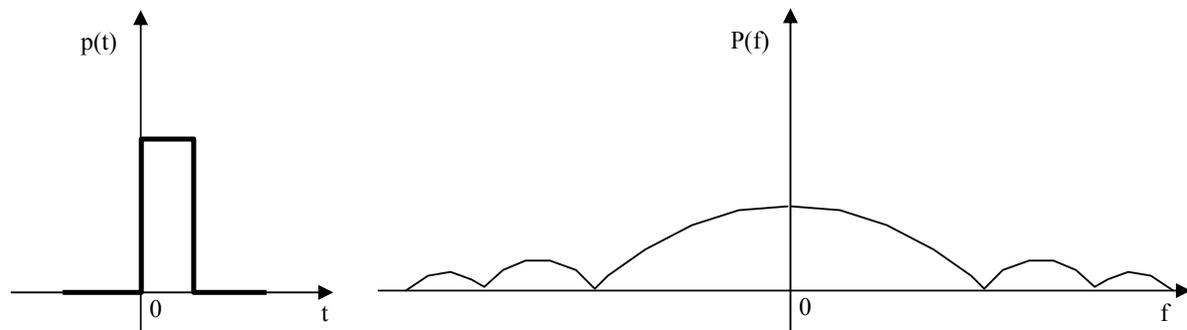


Figura 33

Forma de onda no tempo de um pulso, e espectro de amplitude (na frequência) desse pulso.

No tempo, a amostragem ideal da Figura 30 ficaria da forma da Figura 34. Já não é um ponto, mas um patamar constante, com o valor amostrado, de duração igual à do pulso.

Em termos de frequência, vai haver um processo análogo ao da Figura 32. No entanto, a envolvente dos vários espectros já não é uma recta (a pontado na Figura 32), mas a forma da função da parte direita da Figura 33. Este processo está representado na Figura 35. Como se pode ver, embora a figura não o mostre muito bem, vai existir uma deformação dos espectros, ocasionando uma distorção do sinal no final que terá de ser compensada. A Figura 35 está muito exagerada para se perceber o efeito da amostragem por pulsos. A situação não é tão grave para o espectro mais baixo (a cinzento) que é o que nos interessa, pois a curva do lóbulo principal nesse ponto assemelha-se bastante à recta. Mas de qualquer modo, existe sempre uma deformação.

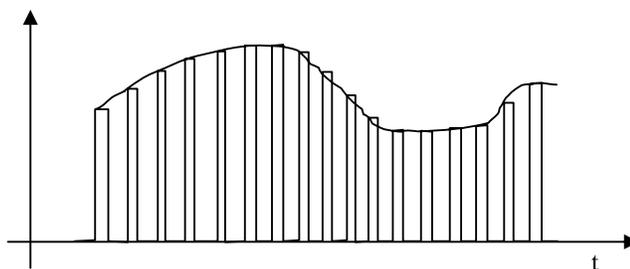


Figura 34
Amostragem de um sinal $m(t)$, por pulsos

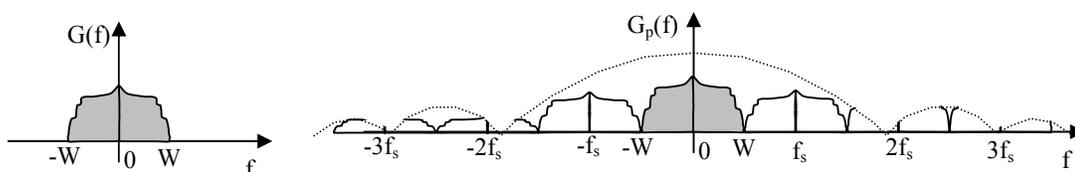


Figura 35
Espectro do sinal original contínuo limitado na frequência a W Hz, e espectro da sua versão amostrada com pulsos a um ritmo de amostragem igual a $2W$ Hz.

1.8.3.2. Quantização

Voltemos outra vez à amostragem por impulsos unitários. A sequência $\{m(n/2W)\}$, embora seja já discreta no tempo, é contínua nos valores de amplitude que pode tomar. Não é ainda uma sequência digital. O processo para que possa tomar apenas valores discretos de amplitude designa-se por **quantização**. A ideia é aproximar o valor de cada amostra contínua a um valor de um conjunto finito de valores de tal modo que se minimize o erro que inevitavelmente se comete. A Figura 36 mostra como a operação é feita para o sinal discreto da Figura 30, ampliado um pouco para ficar mais visível. O valor que se escolhe é o valor médio de cada intervalo que contém a amostra.

O assunto dos “intervalos” não é completamente trivial. Primeiro, toda a gama de possíveis valores que o sinal pode ter tem de estar abrangida por intervalos. Isto é, tem de haver intervalos nos valores mínimos que o sinal pode tomar, assim como nos valores máximos. À relação entre a potência máxima que o sinal pode apresentar e a potência mínima chama-se **gama dinâmica** do sinal. Segundo, se o número total de intervalos for muito grande, cada intervalo é suficientemente pequeno para que o erro que se comete ao aproximar o valor da amostra pelo valor médio do intervalo seja insignificante. Então, o que convinha seria ter muitos intervalos, muito pequeninos, que cobrissem toda a gama dinâmica. Infelizmente, existe uma contrapartida a descrever na operação seguinte, codificação, que faz com que o número total de intervalos não deva ser infinitamente grande. Antes de se passar à próxima operação convém notar que o valor absoluto máximo do erro que se comete, chamado **erro de quantização**, é igual a metade do valor do intervalo. A Figura 37 mostra esse erro.

O erro de quantização **é irrecuperável**, pois o receptor não sabe que valor tinha realmente a amostra, e recebe simplesmente a indicação do intervalo. A partir daí, assume que o valor da amostra é o valor médio do intervalo. Por ser irrecuperável, na construção de um sistema há que dar bastante atenção ao modo como os intervalos são definidos para se distorcer o sinal apenas no montante necessário que alguma contrapartida o exija.

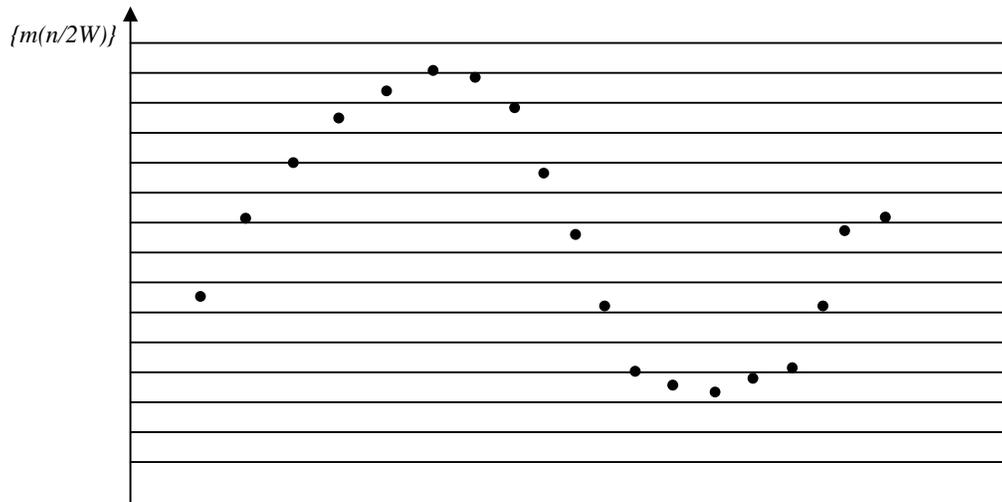


Figura 36
Intervalos de quantização para discretizar amostras na amplitude.

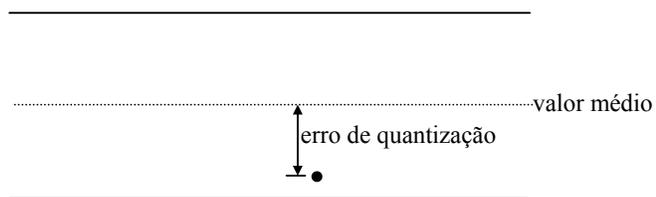


Figura 37
Erro de quantização.

O sinal discreto tanto no tempo como na amplitude é agora um verdadeiro sinal digital. A sua forma é em escada, como mostra a Figura 38. Na figura são visíveis os erros que se estão a cometer relativamente ao sinal original (para melhor se verem esses erros foram unidas as amostras do sinal original).

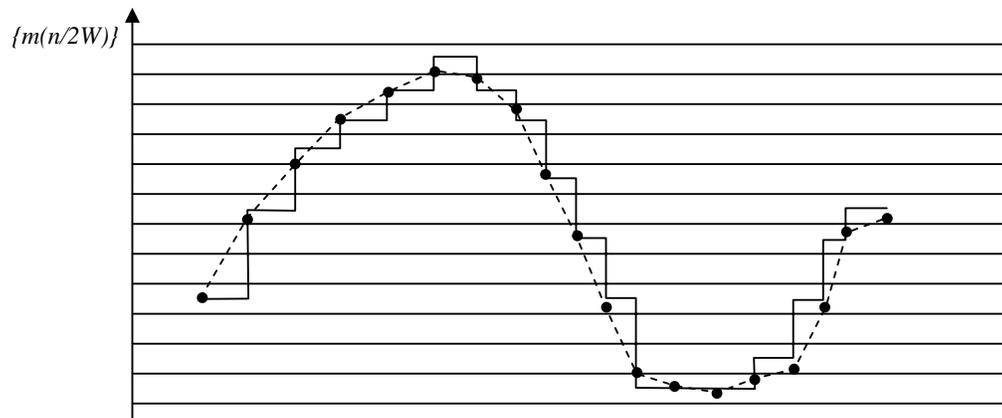


Figura 38
Forma de onda de um sinal amostrado e discretizado na amplitude.

1.8.3.3. Codificação

A operação seguinte consiste em “dizer” ao receptor em que intervalo uma certa amostra está. Primeiro do que tudo essa informação tem de ser “dita” em, pelo menos, o tempo entre amostras ($1/2W$ segundos), pois passado esse tempo é necessário dizer a informação da outra amostra. Agora o problema reside em como se “diz” isso ao receptor. O modo escolhido foi o de numerar os intervalos, simplesmente, começando em 0 até ao número de intervalos menos um. Muito bem. Já temos a identificação por números. Como vamos então transmitir esses números? Um modo (existem mais) é usar uma **codificação** binária. Assim, se houver 256 intervalos, usam-se 8 bits para os identificar. O intervalo 0 pode ser 0000 0000 e o intervalo 200 seria 1100 1000.

Qual a realidade a que se chegou com estas escolhas? Tem de se transmitir 8 símbolos por cada amostra. O tempo máximo para se transmitirem esses símbolos deve ser inferior ou igual a $1/2W$ segundos. Por exemplo, se quisermos aproveitar as frequências de um sinal de voz até aos 4.000 Hz, a frequência de amostragem tem de ser o dobro, portanto, 8.000 Hz. Ora a 8.000 Hz vamos ter uma amostra a cada 125μ segundos. Se cada amostra tiver 8 símbolos então cada símbolo deve durar, no máximo, $15,625 \mu$ segundos. Como estamos a transmitir 8.000 vezes 8 símbolos por segundo, estamos a transmitir 64.000 símbolos por segundo, ou 64 Kbit por segundo⁸. Simples, não é?

Mas como transmitir símbolos para o outro lado? A melhor forma é inventarmos pulsos com uma certa forma para os bits 0 e 1. Na realidade, existe uma miríade de formas de pulsos diferentes que podem ser escolhidas. Durante a disciplina vamos estudar as vantagens de se escolher uma forma de pulsos ou outra. Por agora vai-se apresentar simplesmente dois exemplos diferentes para se ter uma ideia de como a forma de um pulso pode ser. Um conjunto de pulsos muito simples está representado na Figura 39. Para o símbolo 0 escolhe-se simplesmente a ausência de tensão na linha, durante todo o tempo do símbolo. Para o símbolo 1 escolhe-se colocar uma tensão de D Volt na linha durante todo o tempo do símbolo. Na parte de baixo da figura está representada a sequência que se enviaria para uma amostra que estivesse no intervalo 200, usando a numeração explicada atrás (1100 1000).

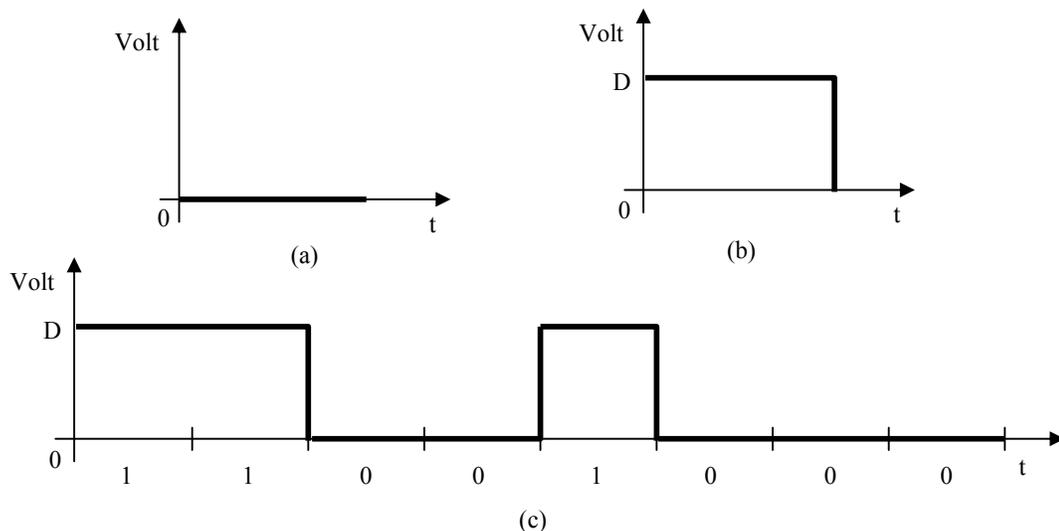


Figura 39

Formas de onda de pulsos e de uma sequência de símbolos (a) símbolo 0, (b) símbolo 1, (c) sequência do intervalo 200.

⁸ Quem já tem algum conhecimento de telefones RDIS sabe que, realmente, este é o ritmo usado para se transmitir conversas telefónicas digitalmente nessa rede.

Outro conjunto de pulsos também muito simples está representado na Figura 40. Para o símbolo 0 escolhe-se colocar uma tensão de $-D$ Volt, durante todo o tempo do símbolo. Para o símbolo 1 escolhe-se colocar uma tensão de D Volt na linha durante todo o tempo do símbolo. Na parte de baixo da figura está novamente representada a sequência de uma amostra no intervalo 200.

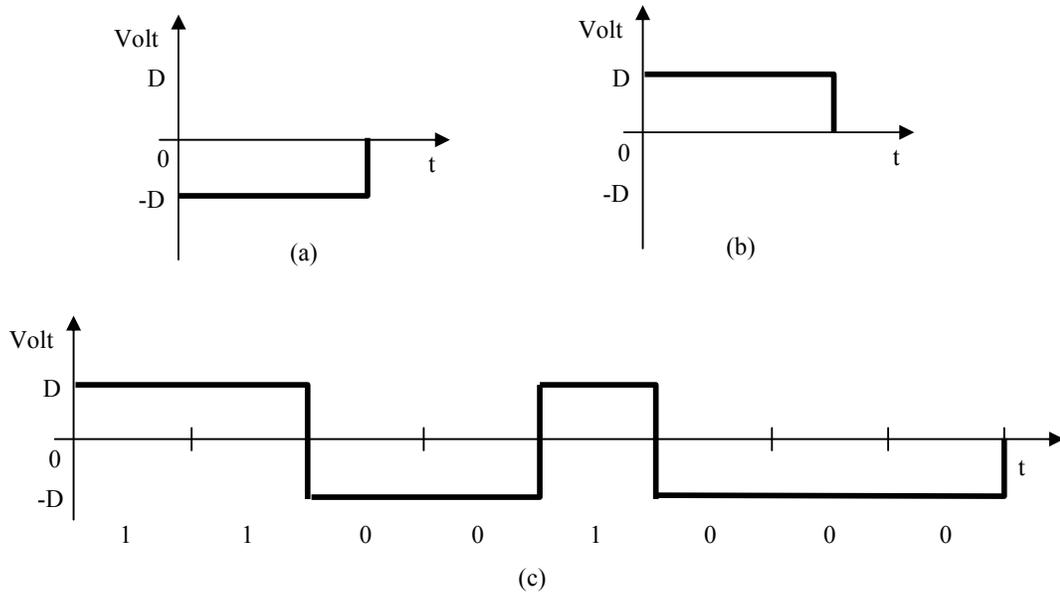


Figura 40

Formas de onda de pulsos e de uma sequência de símbolos (a) símbolo 0, (b) símbolo 1, (c) sequência do intervalo 200.

Percebe-se agora que o número de intervalos não possa ser infinitamente grande. Se houvesse, por exemplo, pouco mais de um milhão de intervalos, para o erro de quantização ser mesmo pequeno, cada intervalo era identificado por 20 bits. Ora a 8.000 Hz de amostragem, em vez de se transmitir a 64 Kbps com amostras de 8 bits, teríamos um ritmo de 160 Kbps pois cada amostra teria 20 bits. Cada conversa desta “alta qualidade” ocupa o “espaço” usado presentemente por duas conversas e meia. Um desperdício de dinheiro para a operadora de telecomunicações. Não faz muito sentido não ganhar mais dinheiro, especialmente se os clientes estão contentes com a qualidade actual dos telefones fixos a 64 Kbps.

... E a nível espectral? Qual a ocupação agora de um sinal modulado que tem frequências até W Hz? O estudo a nível espectral complica-se um pouco. Primeiro deixámos de ter sinusóides, tão fáceis de compreender, para ter pulsos. Uma sinusóide tinha uma frequência pura. Já foi visto que o espectro de um pulso é da forma da função representada na Figura 33. Tudo fica um pouco mais complicado pois a forma de onda do sinal modulado tem a ver com a sequência de 1s e 0s que são de facto enviados para a linha, e isso depende da numeração dos intervalos e em que intervalo o sinal está no momento da amostragem. Claramente, usando os tipos de pulsos anteriores, a pior situação (maiores frequências) é quando existe precisamente 0101 0101, e as melhores situações (menores frequências) são para os intervalos 0000 0000 e 1111 1111, pois geram uma tensão constante na linha que se continuasse até ao infinito teria frequência zero. Normalmente, no estudo espectral de um sinal modulado em PCM escolhe-se a pior situação possível para ver se não é muito distorcido pelo meio de comunicação.

Na pior situação, naquela em que existe uma sequência alternada de zeros e uns, tem-se uma sequência de pulsos em que cada um é como o da Figura 33. O espectro dessa sequência é parecido com o espectro da parte direita da Figura 33, e está esboçado na Figura 41. Existem formas de pulsos que ocupam mais frequências do que outras. O uso de pulsos mais “esbanjadores” só faz sentido se as características deles

forem tais que tragam vantagens nalguns aspectos, como por exemplo um melhor comportamento relativamente ao ruído (uma maior imunidade ao ruído), uma maior facilidade de sincronização de relógios, etc..

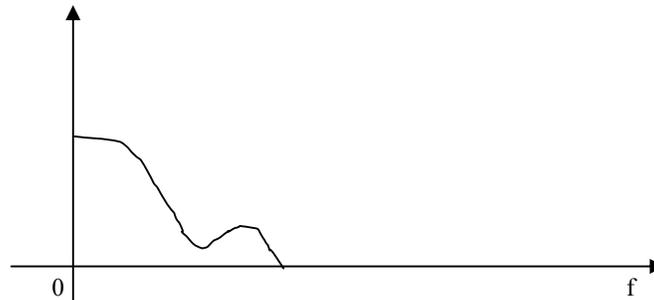


Figura 41
Espectro hipotético de um sinal modulado em PCM.

Vê-se também na Figura 41 que a ocupação na frequência é diferente da dos outros métodos de modulação. No PCM **não existe** uma “passagem” para grandes frequências, mas sim logo o uso das frequências iniciais. A sistemas como este chamam-se de **banda de base**. Os sinais normais de voz, de 400 a 4.000 Hz chamam-se também sinais em banda de base. Às ondas moduladas em AM, FM, etc., chamam-se **sinais passa-banda**, pois ocupam apenas uma faixa de frequência muito longe do zero Hertz.

Um último aspecto para pensar nesta secção do PCM prende-se com a relação entre o espectro do sinal original e o espectro do sinal modulado. Veja-se o que acontece na modulação de amplitude. Se o sinal de banda de base for estreito (até uma frequência de W Hz), ele é transladado e o seu espectro em altas frequências é estreito. Se o sinal de banda de base for largo (até uma frequência Y maior do que W Hz), o seu espectro modulado é também mais largo do que no caso anterior. Pelo que foi dito até agora para a modulação por pulsos, parece que qualquer que seja a ocupação de frequência de um sinal, o espectro do sinal modulado tem a ver com os pulsos e não com o sinal propriamente dito. Um sinal largo, depois de modulado, ocuparia a mesma frequência do que um sinal estreito modulado, pois a ocupação espectral tem a ver com os pulsos. Algo não faz sentido!...

Antes de explicar este aparente contracenso, vamo-nos focalizar novamente na ocupação espectral dos pulsos. Um pulso muito largo (que dure muito tempo) tem uma certa ocupação espectral que é relativamente pequena. Basta pensar que se ele fosse mesmo muito grande a tendência seria para zero Hertz, pois o sinal não se mexe (estava sempre a D Volt). Por outro lado, se o pulso for mesmo muito curto existem componentes de muito alta frequência, pois temos uma perturbação muito rápida. Já vimos que um pulso assim tão estreito se pareceria com um impulso unitário e a tendência do seu espectro seria para uma recta, ocupando assim uma gama de frequências infinita. A Figura 42 mostra de um modo qualitativo esse fenómeno, apresentando as formas de onda de dois pulsos e as suas hipotéticas ocupações espectrais. Concluindo, se os pulsos forem muito curtos eles vão usar muito mais frequências.

Pensemos agora nas frequências dos sinais a modular. Consideremos dois sinais: um primeiro sinal, A , com pouca ocupação espectral, cuja frequência maior é 100 Hz (o nosso sinal estreito); e um outro sinal, B , com componentes de mais alta frequência cuja frequência maior é 50.000 Hz (o sinal largo). Consideremos também que vai haver o mesmo número de intervalos de quantização para os dois sinais. Podem ser 1024 intervalos para poderem ser codificados com 10 bits, e simplificar os cálculos.

Para modular o sinal A temos de ter uma frequência de amostragem de pelo menos 200 Hz, o que dá um intervalo entre amostras de $1/200 = 5$ mseg, e uma duração de cada pulso de $5/10 = 500$ μ seg.

Para modular o sinal B temos de ter uma frequência de amostragem de pelo menos 100.000 Hz, o que dá um intervalo entre amostras de $1/100.000 = 10 \mu\text{seg}$, e uma duração de cada pulso de $10/10 = 1 \mu\text{seg}$ (pulsos quinhentas vezes mais estreitos do que os pulsos usados para o sinal A!...).

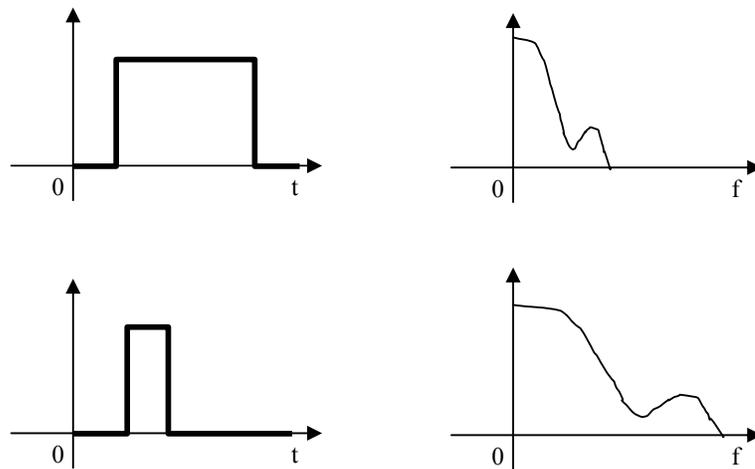


Figura 42
Formas de onda e espectros hipotéticos de dois pulsos.

Conclui-se assim que um sinal com componentes de grandes frequências, ao ser modulado em PCM, vai também ocupar mais frequências do que um sinal com componentes de menor frequência, pois exige pulsos muito mais estreitos.

1.8.4. Modulação por Pulsos Diferencial

A modulação por pulsos diferencial (DPCM) é uma optimização do PCM. A ideia base é que os sinais de voz não mudam de amplitude muito depressa no tempo (os nossos músculos vocais não conseguem fazer isso). Portanto, exagerando, o sinal de voz é mais parecido com a parte (a) da Figura 43 do que com a parte (b). Na figura estão mostrados também os momentos de amostragem, para se perceber a relação.

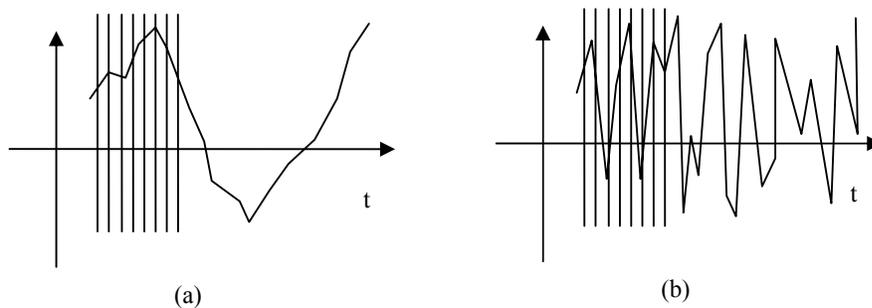


Figura 43
(a) Hipotético sinal de voz humana (exagerado na suavidade); (b) Outro sinal

Ora, se o sinal de voz humana tem esta característica, não vale a pena ter a possibilidade de poder codificar o sinal em qualquer intervalo possível em qualquer momento no tempo. Isto é, se o sinal está muito positivo numa dada altura, ele nunca irá ficar muito negativo na amostra seguinte. Como é que se pode então tirar partido disto?

A ideia é pensar no DPCM como tendo dois passos distintos: a bruxaria e o PCM. Vamos começar a descrição pelo emissor. Na parte de bruxaria o emissor tenta adivinhar o valor da próxima amostra. Na altura dessa amostra é verificado o seu valor real e medido o erro entre o que se adivinhou e o valor real. Esse erro é o sinal que vai ser transmitido para o receptor usando o PCM. Se o passo de bruxaria for muito bom, o emissor adivinha muito bem o valor e o erro que se transmite é sempre muito pequeno ou mesmo zero muitas vezes. No receptor existe também **exactamente o mesmo método** de adivinhar (isto é importantíssimo). O receptor adivinha o valor da próxima amostra e depois recebe o erro do emissor e corrige a sua adivinha pelo valor real que a amostra tem. Imaginemos então que temos 1024 intervalos possíveis (para dar os tais 10 bits e ser fácil fazer contas). A Figura 44 mostra na parte direita o uso desses intervalos no PCM normal. Na parte esquerda está mostrado o erro que se obteve em cada amostra com um certo circuito de adivinhar. Estes valores foram inventados para a figura. Como se pode ver os intervalos são muito mais pequenos e consegue-se com o mesmo número de intervalos uma qualidade muito maior. Poder-se-ia ter a mesma qualidade do PCM, mas com menos intervalos (e menos bits, obviamente). Existe um standard internacional que consegue manter a actual qualidade de voz da rede telefónica com metade da largura de banda (32 kbps em vez de 64 kbps).

O conceito do DPCM é assim muito simples: é um PCM aplicado ao erro que se comete ao adivinhar a amostra.

Um problema grave do DPCM é se o valor da adivinha (o erro) é muito grande. Passa a gama dinâmica que se definiu para o erro e faz com que o emissor e o receptor se dispersem pois tentam adivinhar amostras futuras a partir de momentos presentes diferentes... Pense no que acontece se o erro real for dado pelo ponto A, mas o receptor recebe o ponto B, evidentemente.

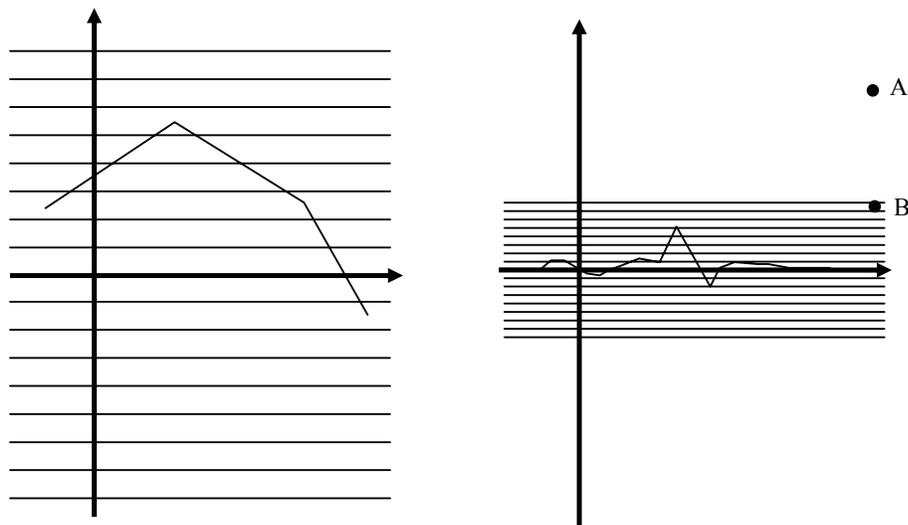


Figura 44

Intervalos no caso de PCM e de DPCM

Mas isto de prever o futuro nunca deu muito bom resultado...

Na década de 1960 começaram a usar-se modems para ligar computadores usando a rede telefónica. Ora os modems não têm músculos vocais e conseguem produzir sinais que têm variações muito abruptas no tempo. De repente, este modo de modular o sinal, modulação por pulsos diferencial, teve de ser encarado com muita cautela na rede telefónica pois saturava muitas vezes (a solução foi ter dois algoritmos diferentes para adivinhar as amostras – um dedicado à voz e outro aos dados)...

1.8.5. Modulação Delta

A modulação Delta é outro tipo de optimização. A ideia está ilustrada na Figura 45, e consiste em ter apenas dois intervalos: um para cima e outro para baixo do sinal no momento. Assim, se na amostra actual o sinal subiu relativamente à amostra anterior (ou mais exactamente, está acima do valor de referência da modulação Delta no momento) envia-se o código do intervalo de cima (por exemplo, 0). Se desceu, envia-se o código do intervalo de baixo (por exemplo, 1). Consegue-se, deste modo, enviar **UM** bit por amostra. É claro que o sinal tem de ser ainda mais suave do que no caso da modulação por pulsos diferencial, pois só existe um intervalo. Um modo de conseguir isso é aumentar o ritmo de amostragem muito para além do dobro da frequência máxima do sinal. A Figura 45 mostra o sinal original, a sequência binária que vai na linha e o sinal de referência da modulação Delta (que é gerado tanto no emissor, como no receptor). O modo mais fácil de pensar é imaginar que o receptor aumenta o valor da amostra presente com um certo valor, ou diminui com o mesmo valor consoante receba 0 ou 1, respectivamente.

A modulação Delta tem outras boas características e outras tantas limitações que serão estudadas na disciplina.

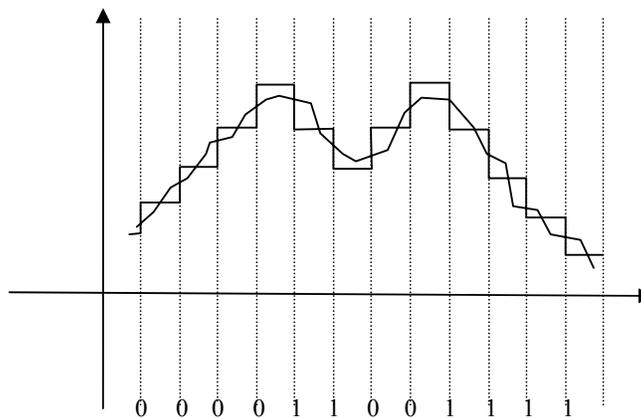


Figura 45
Modulação Delta: sinal original; sinal de referência; e sequência binária

1.9. Técnicas de Modulação Avançadas

No ponto anterior, considerou-se sempre modulações binárias e simples (alteração de apenas um parâmetro do conjunto amplitude, frequência e fase). Neste ponto vão-se descrever modulações com um factor maior do que binárias (cada estado/símbolo corresponde a mais do que um bit) e a alterações de mais de um parâmetro de cada vez.

1.9.1. QPSK – QuadriPhase-Shift Keying

A modulação PSK tem dois valores de diferença de fase – por exemplo 0° e 180° . A Figura 46 mostra um diagrama, chamado de **diagrama de constelação**, em que os pontos indicam os valores legais para a fase. Como se vê existem dois valores legais – 0° e 180° .

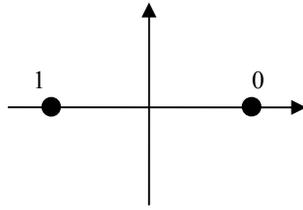


Figura 46

Diagrama de constelação do PSK

Podíamos agora pensar em ter quatro valores legais, em vez de dois. Isto quer dizer que a onda pode estar a ser enviada com um de quatro atrasos possíveis durante um certo tempo. Neste caso, com quatro possibilidades, cada valor vai significar dois bits e não apenas um. Este tempo em que a onda se mantém com o mesmo atraso, que anteriormente designámos por tempo de bit, deve agora ser chamado de **tempo de símbolo**, e cada símbolo representa dois bits. A esta modulação chama-se, QPSK (*QuadriPhase Shift Keying*). Duas possibilidades para valores legais da fase estão mostrados na Figura 47, com os valores dos respectivos bits que representam.

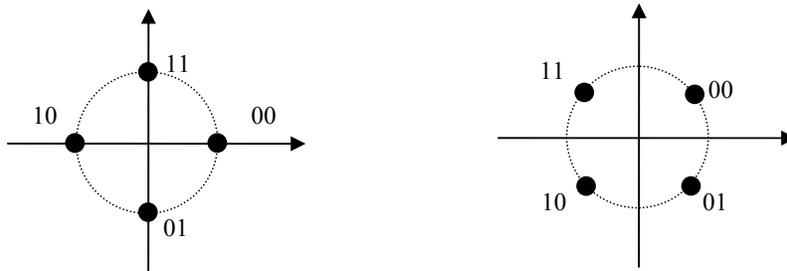


Figura 47

Dois diagramas possíveis para QPSK

1.9.2. M-PSK

Já que estamos com a mão na massa, porque é que não definimos oito valores possíveis para a fase? Cada valor irá representar três bits! ... E porque não 16 valores para a fase? Cada valor vai representar quatro bits! A Figura 48 mostra um diagrama de constelação de um 16-PSK. Repare que os pontos estão todos sobre uma circunferência. Isso significa que a amplitude máxima do coseno é sempre a mesma (o raio da circunferência) e apenas se muda o atraso (o ângulo com o eixo das abcissas).

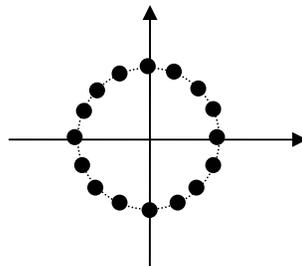


Figura 48

Diagrama de constelação de 16-PSK

1.9.3. QAM revisitado

Um problema com a Figura 48 é que os pontos vão ficando todos “em cima” uns dos outros. Tente imaginar um diagrama de constelação para 256-PSK, em que cada valor representa oito bits. A consequência é que se houver uma má decisão do atraso de fase no receptor, dado que são todas tão próximas, pode-se decidir que o que se recebeu são outros oito bits que não os enviados. Se logo por azar, a sequência de bits for toda diferente de um ponto para outro, obteve-se oito bits errados de uma só vez. Já agora o modo de atribuir bits aos valores de atraso de fase pode minimizar este problema. Tente pensar como resolveria este assunto, mas vamos estudá-lo na disciplina.

Como é que se pode resolver então este problema dos pontos estarem tão “juntinhos”? Uma maneira simples é poder variar também a amplitude da onda. A Figura 49 mostra um diagrama de constelação para 16 valores possíveis. Como se vê, ao permitir que os valores saiam de cima da circunferência conseguiu-se uma separação maior entre eles. Normalmente, este tipo de modulação é mais simples de se pensar como se se fizesse uso da técnica de QAM estudada em cima. Isto é, imagina-se que se estão a enviar duas ondas – um seno e um coseno. Nesta altura da disciplina ainda é complicado perceber porquê e fica aqui apenas a indicação para justificar porque é que a este modo de modulação se chama 16-QAM (ou M-QAM para se ser mais geral).

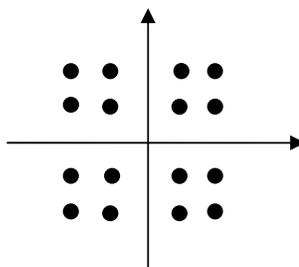


Figura 49

Diagrama de constelação para 16-QAM

Ao contrário das modulações M-PSK, as modulações M-QAM não têm envolvente constante. Isto é, nas M-PSK o coseno tem sempre a mesma amplitude máxima, enquanto que nas M-QAM a onda pode ter várias amplitudes. Para a Figura 49 tente ver quantas amplitudes diferentes pode ter a onda...

1.9.4. MSK – Minimum Shift Keying

A modulação mais complicada que iremos estudar na disciplina chama-se MSK (*Minimum Shift Keying*). Uma variação muito próxima dela é usada na geração GSM dos telemóveis. Mas a sua explicação é simples. A transmissão é feita com dois bits (não é binária, portanto). Nas modulações PSK e QAM anteriores, a fase mantinha-se constante durante todo o tempo de símbolo. Na MSK isso não acontece. A duração de cada símbolo é dividida em duas partes. Vamos sempre considerar que no início da primeira parte o atraso da onda é zero. Se o primeiro bit que queremos transmitir for 0, a fase passa de um atraso de 0° para um atraso de -90° desde o momento inicial até ao fim da primeira parte do símbolo. Esta variação é linear. Se o primeiro bit for 1 então a fase vai variar desde um atraso de 0° até um atraso de 90° . Assim ao fim da primeira parte temos uma onda com um atraso de -90° ou de 90° . Note que ao longo da duração dessa parte o atraso está sempre a mudar. Para o segundo bit as coisas acontecem do mesmo modo. Isto é, se o segundo bit for 0 a onda vai atrasar-se outra vez de -90° (Se estava em -90° passa para -180° . Se estava em 90° passa para 0°). Se for 1 atrasa-se de 90° (Se estava em -90° passa para 0° . Se estava em 90° passa para -180°). Parece complicado mas a Figura 50 mostra como é simples. Na parte de cima estão ilustrados os valores de fase em QPSK. Em baixo estão ilustrados os valores de fase

Imaginemos que se liga o computador ao modem pela porta série. O computador envia bits para o modem com a forma de pulsos como os da Figura 39. Ora, em frequência estes pulsos ocupam uma faixa desde 0 Hz até um certo valor. Identificam-se dois problemas. O primeiro é que um canal telefónico só começa em 400 Hz, portanto o sinal não pode ser introduzido directamente na linha. O segundo é que se o ritmo de bits do computador for muito elevado, os pulsos serão muito estreitos e a frequência máxima passará os 3400 Hz do canal telefónico. Assim, o modem vai ter duas tarefas: mudar o sinal para uma frequência mais central do canal telefónico e usar um tipo de modulação que seja menos “esbanjador” de frequência do que os pulsos, para permitir ritmos maiores não “gastando” tanta frequência. O sinal à saída do modem (para a linha telefónica) é um sinal analógico, e não um sinal digital como o da comunicação entre o computador e o modem. Não existem modems tão simples como um de ASK, mas pense que sim. Repare que o sinal ASK é um sinal analógico. O que ele representa é que é um sinal digital. Já agora os modems mais recentes usam 14 bits por símbolo...

Assim, na vida real já vimos dois usos muito importantes da modulação!

1.11. Multiplexagem

A última secção desta Introdução vai abordar o assunto de multiplexagem nos meios de comunicação. A Figura 52 mostra de um modo muito simplificado o comportamento na frequência de um meio de transmissão metálico (cobre, por exemplo)⁹. Existe uma região em que todas as frequências do sinal são propagadas com quase a mesma atenuação. A seguir existe uma região em que as frequências vão sendo cada vez mais atenuadas à medida que a frequência sobe. É já uma região de difícil utilização em Telecomunicações. Finalmente, existe outra região onde a atenuação é tanta que se torna impossível a sua utilização para as Telecomunicações. Para concretizar os conceitos de multiplexagem foi desenhado também o espectro de um sinal de voz, para se ter uma percepção das relações em causa (mais uma vez, interessa mais uma análise qualitativa e não tanto qualquer exercício quantitativo). As escalas do eixo de amplitude para o comportamento do canal e para o espectro de amplitude são também diferentes.

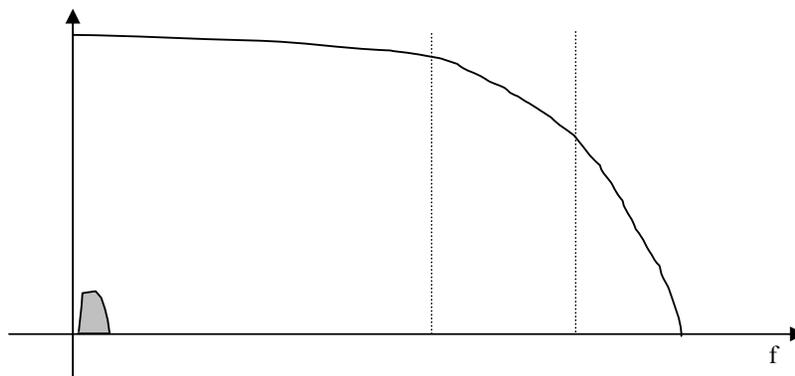


Figura 52

Comportamento na frequência de um meio de transmissão metálico (cobre, por exemplo).

Uma constatação da Figura 52 é que existe ainda muito “espaço” de frequência que se poderia usar no cabo metálico para além do que está a ser usado para apenas um canal de voz. Como pode, então ser esse “espaço” utilizado? Vamos abordar dois modos de o fazer. Na realidade são três os modos mais

⁹ Se fosse o espaço livre, ou uma fibra óptica, a parte que propagava bem o sinal estaria colocada em frequências maiores e não logo a partir do zero, como já deve ser óbvio para o leitor, depois das explicações dadas neste capítulo.

populares, mas o terceiro ficará para mais tarde no curso de Engenharia. O primeiro modo designa-se por multiplexagem por divisão na frequência – FDM (*Frequency Division Multiplexing*), e o segundo por multiplexagem por divisão no tempo – TDM (*Time Division Multiplexing*).

1.11.1. Multiplexagem por Divisão na Frequência – FDM

Na multiplexagem por divisão na frequência a ideia é usar ondas portadoras com diferentes frequências de tal modo que coloquem as conversas uma a seguir à outra no espaço útil de frequências do meio de transmissão. O afastamento entre as diversas frequências das portadoras deve ser tal que o espectro de uma conversa não se sobreponha, nem toque, no espectro da sua vizinha para não provocar interferências. Existe mesmo um espaço de **guarda** para se ficar longe deste tipo de problemas. A Figura 53 mostra o que se acabou de descrever para um sistema de multiplexagem de 8 conversas telefónicas.

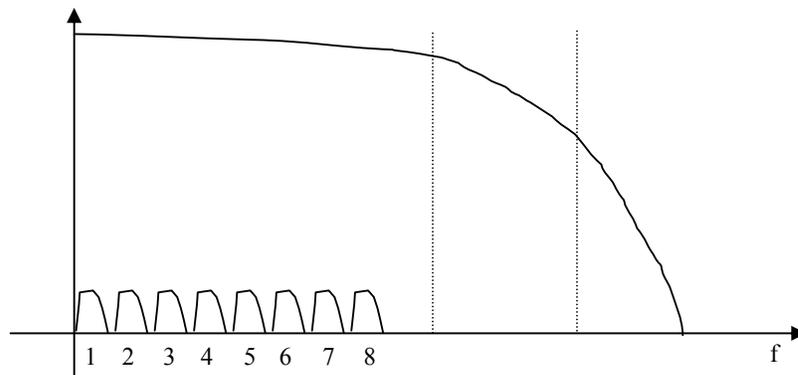


Figura 53
Exemplo da utilização de FDM num meio de transmissão

Com a explosão da rede telefónica, houve a necessidade de colocar cada vez mais chamadas telefónicas no mesmo meio de transmissão. Principalmente quando se ligam centrais grandes umas às outras, como por exemplo uma central em Lisboa e outra no Porto. Assim, foram sendo usados meios com “espaços” de frequência útil cada vez maiores. Passou-se dos pares de cobre para os cabos coaxiais e depois para as fibras ópticas. Hoje em dia existem standards para se colocarem 230.000 conversas telefónicas no mesmo meio de transmissão.

O modo de multiplexagem FDM foi muito usado na rede telefónica analógica e ainda é usado em certos casos de transmissão menos comuns. Hoje em dia, com a digitalização das redes, o modo preferido é a multiplexagem por divisão no tempo que é o tema da próxima subsecção.

1.11.2. Multiplexagem por Divisão no Tempo – TDM

O FDM usa tecnologia electrónica analógica. Com o tempo, esta tecnologia começou a ser desinteressante tanto de um ponto de vista técnico como económico. Começou-se então a usar a multiplexagem por divisão no tempo – TDM. A ideia base é muito simples: imaginemos uma chamada telefónica amostrada a 8.000 Hz com oito bits por amostra. Então, cada amostra tem de ser transmitida a cada 125 μ seg e cada pulso deve durar 15,625 μ seg, como já foi calculado anteriormente. Imaginemos agora que estamos a usar pulsos muito mais curtos do que o necessário. Por exemplo, cada pulso vai durar 0,001 μ seg, ou 1 nano segundo. Se assim for, ao fim de 8 nano segundos (8 nseg) já se transmitiu

esta amostra e temos de esperar 15,617 μseg até transmitir a próxima. A linha está desocupada este tempo todo até ser usada para a próxima amostra.

Então, porque não colocar outra amostra de outra conversa telefônica logo a seguir à primeira? Usar-se-iam mais 8 nseg e o tempo em que a linha ficaria livre seria um pouco menor (já só 15,609 μseg). ... E porque não colocar ainda outra conversa? ...E outra? ...E outra? ...Até preencher todo o tempo livre da linha com o máximo de conversas telefônicas que for possível, até ser necessário outra vez colocar a amostra da primeira conversa?

É assim que funciona o TDM. A linha é utilizada por uma conversa telefônica apenas algum tempo de tempos a tempos. Nos outros intervalos ela é utilizada pelas outras chamadas. Percebe-se também que todo o sistema é muito exigente em termos de sincronização. O receptor tem de saber exactamente o que está a ler para o atribuir à chamada respectiva e, no caso do exemplo, cada pulso dura apenas 1 nano segundo!... A evolução de osciladores (relógios) nas últimas décadas tem sido enorme e tem possibilitado que este modo de multiplexagem seja utilizado com pulsos cada vez menores e com a possibilidade de colocar cada vez mais conversas no mesmo meio. Evidentemente que, quanto menor for o pulso, mais frequência ele ocupa. Assim, tal como FDM, poucas conversas conseguem ser postas em cabos de cobre, mais conversas “cabem” em cabos coaxiais e ainda mais em fibras ópticas. A Figura 54 mostra o esquema geral do TDM com amostras de oito bits, para n conversas.

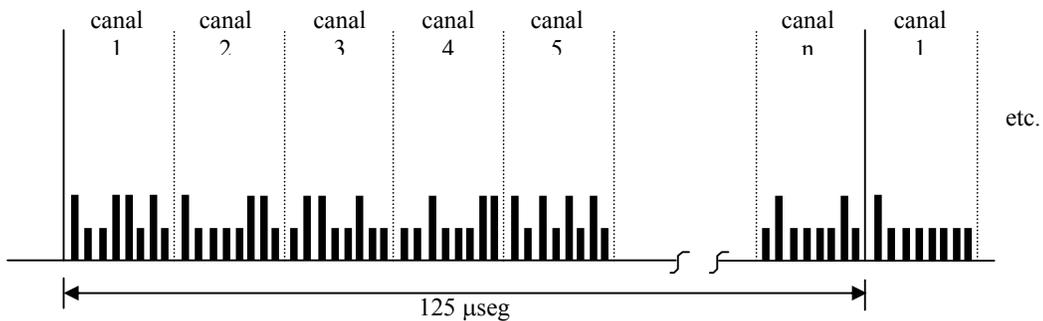


Figura 54
Exemplo da utilização de TDM

1.11.3. QAM revisitado pela última vez

O que se vai escrever neste ponto já deve ser óbvio para o leitor. O QAM, ao conseguir transmitir dois sinais, $m_1(t)$ e $m_2(t)$, ao mesmo tempo é também uma das formas mais básicas de multiplexagem que existem.

Na vida real esta capacidade é mais utilizada para transmitir em paralelo uma sequência de bits do que propriamente para multiplexar duas comunicações independentes. Mas o princípio é o mesmo. Por exemplo, na sequência 0 1 1 1 0 0 1 0 0 os bits ímpares são transmitidos com o coseno e os bits pares são transmitidos com o seno. Assim, a comunicação é feita dois bits a dois bits de cada vez.