

Engenharia de Sistemas e Informática

(Ramo de Sistemas)

(ex - Engenharia de Sistemas e Computação)

Fundamentos de Telecomunicações

w3.ualg.pt/~sjesus/aulas/fdt/

Sérgio M.M. Jesus
sjesus@ualg.pt

Faculdade de Ciências e Tecnologia
Universidade do Algarve
1999/00

Versão 2.0d - 16/Fevereiro/2004

(<ftp://ftp.ualg.pt/users/sjesus/aulas/fdt-V2.0d.pdf> e [fdt-V2.0d-X2.pdf](ftp://ftp.ualg.pt/users/sjesus/aulas/fdt-V2.0d-X2.pdf))

Últimas Revisões (V 2.0, Fevereiro 2004)

O material deste texto de apoio foi inicialmente escrito para a disciplina de Sistemas de Comunicação, do curso de Engenharia de Sistemas e Computação em 1999/00, quando de uma substituição isolada do docente habitual da disciplina. O seu conteúdo foi, nesta segunda versão, adaptado às necessidades da nova disciplina de Fundamentos de Telecomunicações da Licenciatura em Sistemas e Informática tendo em conta, neste primeiro ano, a transição de Engenharia de Sistemas e Computação, que deu origem a um extenso capítulo de revisões que será posteriormente retirado. Os capítulos 3 e 5 são baseados em [1], enquanto os capítulos 4, 6 e 7 se baseiam fortemente em [2]. A parte de aulas práticas inspiram-se em [3]. Do ponto de vista técnico passam a existir: i) uma versão pdf da sebenta com paginação frente e verso, ii) uma outra com duas páginas A5 por cada página A4, mais económica em papel e iii) encontra-se igualmente disponível uma **versão HTML ONLINE completa da sebenta em** w3.ualg.pt/~sjesus/aulas/fdt.

NOTA PRÉVIA

O material contido neste conjunto de apontamentos é cedido a título gratuito e para ser utilizado exclusivamente como texto de apoio à disciplina de Fundamentos de Telecomunicações do curso de Engenharia de Sistemas e Informática da Universidade do Algarve. Este texto poderá (e tem com certeza) erros involuntários, de cujas consequências o autor não poderá ser responsabilizado. A consulta deste texto não dispensa (e aliás aconselha) a de outras obras, nomeadamente as citadas na bibliografia. Boa leitura !...

Índice

1	Introdução	11
2	Revisões	15
2.1	Sinais determinísticos	15
2.1.1	Sistemas lineares e invariantes no tempo (SLIT)	15
2.1.2	Transformada de Fourier	16
2.1.3	Amostragem	18
2.1.4	Energia e potência	24
2.1.5	Sinais passa-banda e a transformada de Hilbert	24
2.1.6	Passagem em banda-base	26
2.2	Sinais aleatórios	27
2.2.1	Noção de processo aleatório	27
2.2.2	Média e variância de um processo aleatório	28
2.2.3	Covariância e correlação	29
2.2.4	Estacionaridade e ergodicidade	31
2.2.5	Energia e potência	32
2.2.6	Filtragem linear de sinais aleatórios	33
2.2.7	Ruído branco	34
2.2.8	Introdução à estimação espectral	36
2.2.9	Sinais cicloestacionários	39
3	Canais de transmissão	40
3.1	Linhas de transmissão	40
3.2	Fibra óptica	42
3.3	Transmissão rádio	44

3.4	Modelos matemáticos de canais de transmissão	45
3.4.1	Canal de ruído aditivo	45
3.4.2	Canal de filtro linear invariante	45
3.4.3	Canal de filtro linear variante no tempo	46
4	Quantidade de informação, quantificação e codificação	47
4.1	Quantidade de informação e entropia	48
4.2	Quantificação	49
4.2.1	Quantificação uniforme	50
4.2.2	Quantificação não uniforme	50
4.3	Codificação	51
4.3.1	Com palavras de comprimento fixo	52
4.3.2	Com palavras de comprimento variável	52
4.3.3	Codificação de sinais analógicos	54
4.3.4	Codificação em sub-bandas	58
4.3.5	Codificação baseada em modelos do sinal	59
5	Modulação analógica de onda sinusoidal	60
5.1	Modulação de amplitude com duas bandas laterais (AM-DSB)	60
5.2	Modulação de amplitude com supressão de portadora (AM-CS)	63
5.3	Modulação de amplitude de banda lateral única (AM-SSB)	63
5.4	Modulação de amplitude em quadratura (QAM)	65
5.5	Modulação de ângulo	66
5.6	Multiplexagem no tempo e na frequência	68
5.7	Desmodulação AM e FM	69
5.7.1	Extracção do envelope complexo	70

5.7.2	Extracção do envelope e da fase	71
6	Modulação digital de sinais	72
6.1	Introdução	72
6.2	Representação de sinais em expansões ortogonais	72
6.3	Modulação por amplitude de impulsos (<i>pulse amplitude modulation - PAM</i>)	76
6.4	Modulação por fase de impulsos (<i>pulse phase modulation - PPM</i>)	78
6.5	Modulação de amplitude em quadratura (<i>Quadrature Amplitude Modulation - QAM</i>)	80
6.6	Modulação de frequência (<i>Frequency Shift Keying - FSK</i>)	81
6.7	Densidade espectral de potência do sinal modulado	82
6.8	Exemplos	84
7	Receptor óptimo em presença de ruído aditivo	87
7.1	O receptor óptimo	87
7.2	O filtro adaptado como receptor óptimo	89
7.3	Em banda passante	92
A	Tabelas e relações particulares	93
A.1	Transformada de Fourier	93
A.2	Relações Trigonométricas Usuais	95
A.3	Desenvolvimentos em série	96
A.4	Algumas relações úteis	97
A.4.1	Integrais	97
A.4.2	Séries	97
A.4.3	Derivadas	97

A.4.4	Trigonometria do círculo	97
B	Séries de Fourier	98
C	Noção de variável aleatória	104
C.1	Probabilidades elementares	104
C.2	Conceito de variável aleatória e de função de distribuição	105
C.3	Função densidade de probabilidade	107
C.4	Esperança matemática	107
C.5	Função característica	108
C.6	Exemplos	109
C.6.1	Distribuição uniforme	109
C.6.2	Distribuição exponencial unilateral	110
C.6.3	Distribuição exponencial bilateral	110
C.6.4	Distribuição normal (ou gaussiana)	110
C.6.5	Distribuição do χ^2 com d graus de liberdade	111
D	Folhas de Exercícios	112
D.1	Revisões sobre sistemas e sinais	112
D.2	Sinais determinísticos I	114
D.3	Sinais determinísticos II	116
D.4	Sinais aleatórios I	118
D.5	Sinais aleatórios II	120
D.6	Sinais aleatórios III	121
D.7	Modulação e desmodulação I	123
D.8	Modulação e desmodulação II	124
D.9	Modulação PAM I	125

D.10 Modulação PAM II	127
E Método de avaliação	130

Calendarização da disciplina

Semana 1

- T** **1. Introdução**
2. Revisões
- T** Sinais determinísticos: sistemas LIT, Transformada de Fourier (sinais contínuos e discretos). Exemplos.
- P** Folha 1: revisões de sistemas e sinais

Semana 2

- T** Sinais determinísticos: amostragem. Energia e potência
- P** Folha 2: sinais determinísticos I

Semana 3

- T** Sinais determinísticos: sinais passa-banda; Transformada de Hilbert; passagem em banda base.
- P** Exercícios de Matlab I

Semana 4

- T** Sinais aleatórios: noção de processo estocástico, média, variância, covariância, correlação; estacionaridade e ergodicidade; Energia e potência. Exemplos
- P** Folha 3: sinais determinísticos II

Semana 5

- T** Filtragem de sinais aleatórios; ruído branco. Noção de estimação espectral. Sinais cicloestacionários. Exemplos.
- P** Exercícios de Matlab II

Semana 6

- T** **3. Canais de Transmissão**
Linhas de transmissão; fibra óptica; transmissão rádio; modelos matemáticos de canais de transmissão.
- P** Folha 4: sinais aleatórios I

Semana 7

- T** 4. **Quantidade de informação,**
quantificação e codificação
Entropia; quantificação; codificação:
Codificação de sinais analógicos: PCM.
- P** Folha 5: sinais aleatórios II

Semana 8

- T** PCM diferencial, PCM diferencial adaptativo
Modulação Delta; codificação em
sub-bandas. Modelos.
- P** Exercícios de Matlab III

Semana 9

- T** 5. **Modulação analógica de onda sinusoidal**
AM-DSB, AM-CS, AM-SSB, QAM; modulação de
ângulo; Modulação no tempo e na frequência
- P** Folha 6: sinais aleatórios III

Semana 10

- T** Desmodulação AM e FM: extração do
envelope complexo e da fase.
Exemplos
- P** Folha 6: Modulação e desmodulação.

Semana 11

- T** 6. **Modulação Digital de sinais**
Introdução; Representação de sinais;
Modulação PAM e PPM (ou PSK)
- P** Folha 7: Modulação e desmodulação.

Semana 12

- T** Modulação QAM e FSK. Representação em
frequência e exemplos.
- P** Exercícios de Matlab IV

Semana 13

- T** **7. Receptor óptimo num canal aditivo**
O receptor óptimo.
- P** Folha 8: Modulação PAM I

Semana 14

- T** O filtro adaptado como receptor óptimo;
O receptor óptimo em banda passante.
Exemplos
- P** Folha 9: Modulação PAM II.

Semana 15

- T** Revisões
- P** N/A

1 Introdução

Sistemas de comunicação são todos os instrumentos que permitem transmitir informação entre dois pontos distantes. Geralmente o ponto de partida da informação é chamado *emissor*, enquanto o ponto de chegada é chamado *receptor*. Entre o emissor e o receptor encontra-se o *meio de transmissão* ou simplesmente *canal* de transmissão. Frequentemente pretende-se que a transmissão da informação seja efectuada nos dois sentidos, através do mesmo canal de transmissão, e então o emissor é também receptor e vice-versa. O esquema de blocos da figura 1.1 representa um sistema de comunicações onde se podem ver os seus três componentes principais, emissor, canal e receptor, mas também os sub-blocos constituintes de cada um destes. Encontramos

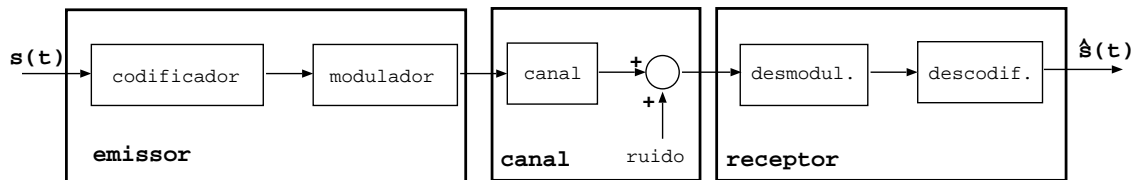


Figura 1.1: esquema de blocos tipo para um sistema de comunicações.

sistemas de comunicação seja no emissor, seja no receptor. Geralmente o papel do sistema de comunicação no emissor é o de adaptar o sinal a transmitir ao canal de transmissão. No receptor trata-se de recolocar o sinal recebido sob a forma original enviada pelo emissor. Tendo em conta que o sinal foi transformado no emissor e sofreu alterações durante a sua passagem pelo canal de transmissão - nomeadamente a influência de ruído e outras distorções - o papel do sistema de recepção será de colocar o sinal recebido numa forma o mais fiel quanto possível em relação ao sinal emitido.

Assim, ao nível do emissor temos pelo menos dois sub-blocos que são o codificador e o modulador. O codificador permite transformar o sinal a transmitir, fornecido pela fonte, de forma a reduzir tanto quanto possível a quantidade de informação necessária e suficiente para permitir a sua recuperação, com uma taxa de erro aceitável, no receptor. Na prática trata-se frequentemente de uma amostragem de sinal analógico em discreto, numa quantificação de amplitude, numa codificação e finalmente numa compressão da informação. Todas estas etapas, algumas das quais são opcionais, dependem por um lado do tipo de sinal que se apresenta à entrada, do sinal que deve entrar no modulador e também do canal de transmissão considerado. O modulador pode ser de muitos tipos e depende essencialmente do canal de transmissão considerado. Em geral podemos classificar o modulador como podendo ser analógico ou digital, consoante o sinal que se apresenta à sua entrada, sendo que o sinal de saída a emitir para o meio físico de transmissão é quase sempre analógico.

O canal de transmissão representa o meio físico onde o sinal é enviado entre o

emissor e o receptor. Como exemplos de canais de transmissão podemos citar: a atmosfera, através da qual podemos enviar ondas electromagnéticas (p. ex. rádio, televisão, telemóvel, etc...), o cabo eléctrico bifilar (telefone, computador a baixo débito), o cabo coaxial (computador, TV por cabo,...) e a fibra óptica, a água (sonar, telefone submarino, etc...) e o vácuo (satélite, naves espaciais, etc..). Cada um destes exemplos de canais de transmissão tem características próprias que necessitam sistemas de comunicação expressamente dedicados. Existe hoje em dia uma grande variedade de sistemas de comunicação, cada um mais sofisticado que o outro. Uma comunicação por telefone pode, por exemplo, utilizar vários canais de transmissão, passando de cabo bifilar para linha digital, depois fibra óptica em seguida via rádio para um satélite, etc... No entanto uma das características essenciais de cada canal de transmissão é a sua capacidade, *i.e.*, em termos simples, a quantidade máxima de informação que podemos transmitir através dele sem distorção apreciável. As noções de quantidade de informação e de distorção serão introduzidas de forma precisa mais adiante.

No receptor efectuam-se as operações inversas daquelas efectuadas no emissor: desmodulação e descodificação. Estas operações dependem fortemente do que foi feito no emissor, considerando no entanto que o sinal de entrada no receptor se encontra transformado pelo canal e, em particular, corrompido com ruído. O sinal recebido tem portanto uma componente aleatória mais ou menos importante dependendo da natureza e nível de ruído do canal. O receptor é normalmente visto como um detector, cuja característica mais importante é a probabilidade de deteção para uma determinada taxa de erro admissível (*receiver operating characteristics - ROC*). Segundo os casos, o desmodulador pode por sua vez ser separado em vários sub-blocos como por exemplo, o filtro de recepção, destinado a compensar a distorção introduzida pelo meio físico de transmissão, o amostrador e o decisor. À saída do receptor dever-se-à obter um sinal $\hat{s}(t)$, tanto quanto possível, idêntico ao sinal emitido.

Exemplo: o eco

A título de exemplo das dificuldades que podem ser encontradas na transmissão de informação entre dois pontos escolhemos o caso da voz num local com eco. Vamos supôr que um explorador tenta comunicar com um companheiro utilizando a voz e que ambos se encontram na escalada de uma ravina rochosa num vale estreito. Sabemos que nesta situação o sinal acústico emitido tem tendência a chegar ao receptor através de mais do que um caminho, *i.e.*, para além do som ouvido em linha recta ouvem-se também repetições desse mesmo som resultado da reflexão nas paredes rochosas. Para simplificar vamos supor que o a um som emitido pelo explorador, o seu companheiro recebe dois: um em linha recta atenuado de uma factor α_0 e atrasado de τ_0 segundos e um outro atenuado de α_1 e atrasado de τ_1 . Assim podemos escrever que para o sinal emitido $s(t)$ obtemos o sinal recebido $r(t)$ tal que

$$r(t) = \alpha_0 s(t - \tau_0) + \alpha_1 s(t - \tau_1) \quad (1-0.1)$$

onde obviamente α_0 e α_1 são < 1 e $\tau_0 = d/c$ onde d é a distância em metros entre os dois exploradores e c é a velocidade de propagação do som no ar. Claro que os coeficientes α_i dependem das distâncias percorridas, mas geralmente, dado que a distância percorrida pelo eco é maior do que a distância percorrida pelo sinal em linha recta temos que $\alpha_1 < \alpha_0$ e que $\tau_1 > \tau_0$. Não sabendo onde é que o som se vai reflectir, torna-se impossível determinar seja α_1 , seja τ_1 . Em particular, se a reflexão se der numa parede próxima da linha de propagação do som em linha recta, pode acontecer que $\Delta\tau = \tau_1 - \tau_0$ seja inferior à duração do sinal $s(t)$. Nessas condições o eco sobrepõe-se ao sinal em linha recta e torna-se difícil, e por vezes impossível, compreender a mensagem enviada. Dizemos neste caso que temos distorção devida a múltiplos caminhos entre o emissor e o receptor. O problema que se coloca no receptor é o de tentar compreender a mensagem. Existem várias formas de resolver o problema. A primeira constatação que se pode fazer é de que se os atrasos τ_0 e τ_1 , ou mesmo apenas a sua diferença $\Delta\tau = \tau_0 - \tau_1$, fosse conhecidos, então poderíamos compensá-los no sinal recebido através de um filtro. Pode-se determinar a função de transferência desse filtro através da TF de (1-0.1)

$$\begin{aligned} R(\omega) &= \text{TF}[r(t)] \\ &= \alpha_0 S(\omega)e^{-j\omega\tau_0} + \alpha_1 S(\omega)e^{-j\omega\tau_1} & (1-0.2) \\ &= H(\omega)S(\omega) & (1-0.3) \end{aligned}$$

onde, para se obter $S(\omega)$ a partir de $R(\omega)$ é necessário aplicar um filtro $G(\omega)$

$$G(\omega) = H^{-1}(\omega) = [\alpha_0 e^{-j\omega\tau_0} + \alpha_1 e^{-j\omega\tau_1}]^{-1}. \quad (1-0.4)$$

No entanto, na prática, os atrasos temporais do eco não são conhecidos e $G(\omega)$ não pode ser facilmente calculado através de (1-0.4). Uma das estratégias para estimar os atrasos consiste em emitir um sinal conhecido do receptor, naquilo que é normalmente chamado como um sequência de treino. Pode-se então correlacionar o sinal recebido com o sinal emitido obtendo-se

$$\begin{aligned} y(\tau) &= \int r(t)s(t-\tau)d\tau \\ &= \int [\alpha_0 s(t-\tau_0) + \alpha_1 s(t-\tau_1)]s(t-\tau)d\tau \\ &= \alpha_0 \int s(t-\tau_0)s(t-\tau)d\tau + \alpha_1 \int s(t-\tau_1)s(t-\tau)d\tau \\ &= \alpha_0 y_s(\tau-\tau_0) + \alpha_1 y_s(\tau-\tau_1) & (1-0.5) \end{aligned}$$

onde $y_s(\tau)$ é a função de autocorrelação do sinal emitido $s(t)$. Tendo em conta que a função de autocorrelação é uma função monótona pode-se deduzir que os dois picos de $y(\tau)$ terão lugar para $\tau = \tau_0$ e $\tau = \tau_1$, de onde estes dois valores podem ser directamente estimados. Este processo de emissão de sequências de treino, conhecidas do emissor e do receptor, antes da mensagem propriamente dita é corrente em quase

todos os sistemas submetidos a fenómenos multicaminhos, como é por exemplo o caso dos actual sistema GSM de telefonia móvel. O problema deste sistema é que por vezes os pontos de reflexão são eles mesmo móveis ou o receptor ou emissor também se movem. Quando a geometria do problema muda ao longo do tempo, os atrasos também se alteram e o filtro estimado a um dado instante já não é válido noutro instante. Dizemos nesse caso que temos um canal de transmissão de estrutura variante no tempo. É claro que nesse caso se pode refazer um novo treino, voltando a construir um filtro correcto. Obviamente que este procedimento não pode vir a ser repetido constantemente pois enquanto o sistema efectua o treino não envia mensagens e a taxa de transmissão do sistema diminui. Existem outras técnicas de extração do sinal útil sem fazer apelo a sequências de treino que utilizam propriedades estatísticas do sinal emitido, para estimar constantemente a estrutura do canal de transmissão. Estas são ditas as técnicas "cegas".

2 Revisões

2.1 Sinais determinísticos

Dissemos no capítulo anterior que a mensagem era transmitida da fonte para o receptor através do canal de transmissão graças a um sistema emissor do lado da fonte e a um sistema receptor do lado do receptor. A informação a transmitir é codificada na fonte sob a forma de um sinal que é geralmente uma voltagem. A variação dessa voltagem ao longo do tempo contém a informação desejada - diz-se que o sinal serve de suporte da informação.

Geralmente, a tensão variável emitida na fonte é uma tensão, dita, analógica, *i.e.*, o sinal pode ser representado por uma função de variável real do tempo, $v(t)$. No entanto, grande parte dos sistemas de comunicação hoje em dia efectuam operações em sinais sob forma discreta e por isso seremos levados a manipular seja sinais contínuos seja discretos e começaremos por rever a noção de amostragem.

2.1.1 Sistemas lineares e invariantes no tempo (SLIT)

Um sistema é considerado linear se obedece às condições de homogeneidade e de sobreposição, tais que

$$x_1(t) \Rightarrow y_1(t) = L[x_1(t)], \quad (2-1.1)$$

$$x_2(t) \Rightarrow y_2(t) = L[x_2(t)], \quad (2-1.2)$$

então

$$k_1x_1(t) + k_2x_2(t) \Rightarrow y(t) = L[k_1x_1(t) + k_2x_2(t)] \quad (2-1.3)$$

tal que $y(t) = y_1(t) + y_2(t)$, com k_1, k_2 reais $\neq 0$. Ainda uma sub-classe dos sistemas lineares são os Sistemas Lineares Invariantes no Tempo (SLIT), para os quais se

$$x(t) \Rightarrow y(t), \quad (2-1.4)$$

então

$$x_0(t) = x(t - t_0) \Rightarrow y_0(t) = L[x_0(t)] = y(t - t_0). \quad (2-1.5)$$

Uma das características essenciais resultantes da definição acima é de que a resposta de um SLIT a uma excitação $x(t)$ é dada pela equação de convolução

$$y(t) = h(t) * x(t) = \int_{-\infty}^{+\infty} h(\tau)x(t - \tau)d\tau \quad (2-1.6)$$

onde $h(t)$ é a resposta impulsiva do sistema, *i.e.*, é a resposta do sistema a um impulso de Dirac $\delta(t)$.

Para sistemas caracterizados no espaço de tempo discreto a equação de convolução escreve-se

$$y(k) = h(k) * x(k) = \sum_{l=-\infty}^{\infty} h(l)x(k-l) \quad (2-1.7)$$

2.1.2 Transformada de Fourier

O integral de Fourier, ou simplesmente transformada de Fourier (TF), que permite passar de um sinal temporal contínuo $s(t)$ (periódico¹ ou aperiódico) a um sinal no domínio da frequência $S(\omega)$ é definida por

$$S(\omega) = \int_{-\infty}^{+\infty} s(t)e^{-j\omega t} dt. \quad (2-1.8)$$

Em geral a $S(\omega)$ chama-se espectro de $s(t)$, mesmo se essa designação deva ser reservada para $|S(\omega)|$, que é também chamada densidade espectral em amplitude. De modo análogo a transformada inversa é definida por

$$s(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} S(\omega)e^{j\omega t} d\omega, \quad (2-1.9)$$

que permite determinar a função temporal $s(t)$ de modo único a partir do seu espectro $S(\omega)$. A transformada de Fourier directa, $S(\omega)$, é geralmente complexa.

TF de um sinal discreto no tempo - DTFT

A transformada de Fourier de um sinal discreto é definida por

$$X(f) = \frac{1}{T} \sum_{k=-\infty}^{\infty} x(k)e^{-j2\pi fkT}, \quad (2-1.10)$$

onde T é o intervalo de amostragem. Do mesmo modo a transformada inversa é obtida por

$$x(k) = \int_{-1/2T}^{1/2T} X(f)e^{j2\pi fkT} df, \quad (2-1.11)$$

no intervalo $[-1/2T, 1/2T]$ visto que $X(f)$ é uma função periódica, de período $1/T$, pois trata-se da transformada de Fourier de um sinal obtido por amostragem. Com efeito esta propriedade demonstra-se facilmente a partir de (2-1.10). Assim

$$X(f + 1/T) = \sum_{k=-\infty}^{\infty} x(k)e^{-j2\pi fkT} e^{-j2\pi k} = X(f), \quad (2-1.12)$$

¹para o caso específico dos sinais periódicos ver apêndice B

visto que $e^{-j2\pi k} = 1$ para qualquer k inteiro. Tendo em mente esta diferença essencial entre a TF de sinais contínuos e de sinais discretos podemos demonstrar todas as outras propriedades da TF para os sinais discretos de modo análogo ao dos sinais contínuos. Frequentemente, para simplificar a notação, as definições (2-1.10) e (2-1.11) são utilizadas com um intervalo de amostragem $T = 1$.

Exemplos:

A) sinal porta: calcular a TF do sinal discreto

$$p(k) = \text{rect}_N(k) = \begin{cases} 1, & -N/2 \leq k \leq N/2 - 1 \\ 0, & \text{para todo outro } k. \end{cases}$$

Aplicando a relação de definição

$$P(f) = \sum_{k=-N/2}^{N/2-1} e^{-j2\pi f k},$$

fazendo $l = k + N/2$, tem-se que $k = l - N/2$ e por substituição

$$P(f) = e^{j\pi f N} \sum_{l=0}^{N-1} e^{-j2\pi f l},$$

o valor do somatório pode ser calculado observando que se trata de uma progressão geométrica de N termos e de razão $r = e^{-j2\pi f}$. Assim

$$P(f) = e^{j\pi f N} \frac{1 - e^{-j2\pi f N}}{1 - e^{-j2\pi f}},$$

que se pode escrever também

$$P(f) = e^{j\pi f} \frac{\sin(\pi f N)}{\sin(\pi f)}.$$

B) sinal amostrado: dado o sinal amostrado, sob a sua forma de modulação de amplitude de impulsos

$$\hat{x}(t) = \sum_k x_k \delta(t - kT)$$

demonstrar que a TF, $\hat{X}(\omega) = \text{TF} [\hat{x}(t)]$ é igual à DTFT da sequência discreta x_k , *i.e.*, que

$$\hat{X}(\omega) = \text{DTFT}[x_k].$$

Com efeito, calculando a TF de $\hat{x}(t)$ temos

$$\hat{X}(\omega) = \int_{-\infty}^{\infty} \hat{x}(t)e^{-j\omega t} dt \quad (2-1.13)$$

$$= \int_{-\infty}^{\infty} \sum_k x_k \delta(t - kT) e^{-j\omega t} dt \quad (2-1.14)$$

$$= \sum_k x_k \int_{-\infty}^{\infty} \delta(t - kT) e^{-j\omega t} dt \quad (2-1.15)$$

$$= \sum_k x_k e^{-j\omega kT} \quad (2-1.16)$$

$$= X(\omega) \quad (2-1.17)$$

$$= \text{DTFT}[x_k]. \quad (2-1.18)$$

No entanto, a TF de sinais discretos - DTFT, é de relativamente pouca utilidade na prática, pois que de modo a permitir o processamento do sinal frequencial por calculadores numéricos, é necessário que a variável f seja também discreta. Assim, definiremos mais tarde a transformada de Fourier discreta que chamaremos de modo abreviado TFD (ou DFT=Discrete Fourier Transform).

2.1.3 Amostragem

Neste capítulo estudaremos as características e as propriedades do processo de amostragem. Este processo, descrito na figura 2.1 consiste em formar, a partir de um sinal contínuo $x(t)$, uma nova função, $\hat{x}(t)$ chamada função amostrada. Esta função obtém-se a partir da função inicial $x(t)$ através de um processo de amostragem periódico (de período τ segundos). Noutras palavras, a função $\hat{x}(t)$ é obtida pelo produto de $x(t)$ com a função de amostragem $s(t)$, que é uma série periódica de impulsos estreitos (em relação a τ). Assim a função $x(t)$ modula em amplitude $s(t)$ para formar $\hat{x}(t)$. A operação inversa consiste no processo de reconstrução do sinal inicial $x(t)$ a partir das amostras da função amostrada $\hat{x}(t)$. Isto é realizado na figura 2.1 por um filtro ideal de interpolação. Consideremos um sinal $x(t)$, passa-baixo, com uma banda limitada,

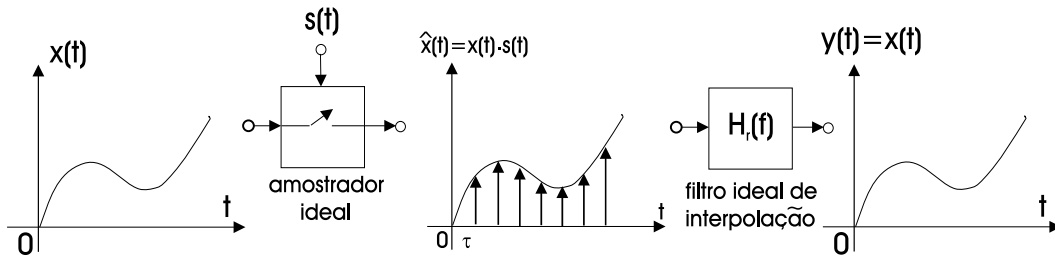


Figura 2.1: processo de amostragem e de reconstrução.

tendo um espectro $|X(f)|$ que é nulo fora de uma banda $(-W, W)$ (ver figura 2.2). Num problema de comunicação servimo-nos em geral de uma portadora de frequência ω_0 , de tal modo que o sinal modulado é

$$v(t) = x(t) \cos \omega_0 t, \quad \omega_0 = 2\pi f_0 \quad (2-1.19)$$

Como sabemos que a representação frequencial de $\cos \omega_0 t$ é constituída por dois Diracs colocados a $\omega = -\omega_0$ e $\omega = \omega_0$ o produto temporal da (2-1.19) torna-se numa convolução no domínio da frequência. O resultado é

$$V(f) = X(f) * \text{TF}[\cos(2\pi f_0 t)] = \frac{1}{2}X(f + f_0) + \frac{1}{2}X(f - f_0) \quad (2-1.20)$$

Isto está ilustrado na figura 2.3. Este resultado pode ser generalizado para o caso

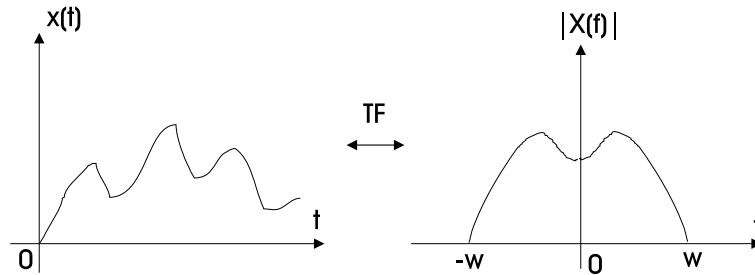


Figura 2.2: espectro do sinal original.

em que $s(t)$ é uma soma de funções periódicas a frequências múltiplas de ω_0 , isto é, $\pm k\omega_0$. Neste caso o produto de (2-1.19) dá no domínio da frequência uma repetição do espectro de $x(t)$ às frequências harmónicas $\pm k\omega_0$.

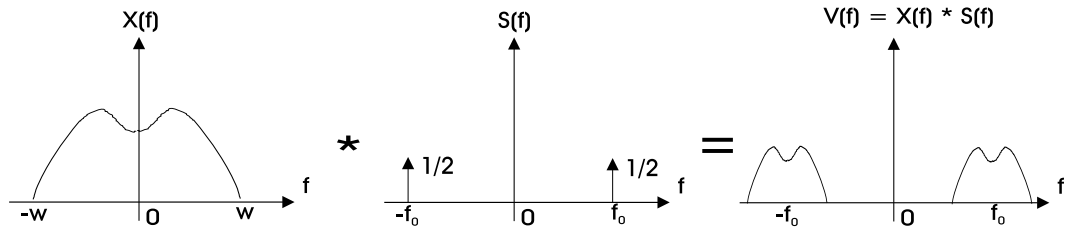


Figura 2.3: espectro do sinal amostrado.

Teorema de amostragem (ou de Shanon)

A forma que deve ter a função periódica $s(t)$, para realizar uma amostragem ideal, é dada por uma série periódica de impulsos de Dirac. Noutras palavras, pode-se definir a função de amostragem ideal por

$$s(t) = \sum_{n=-\infty}^{\infty} \delta(t - n\tau), \quad (2-1.21)$$

que evidentemente tem como espectro

$$S(f) = \frac{1}{\tau} \sum_{k=-\infty}^{\infty} \delta(f - kf_s), \quad (2-1.22)$$

onde $f_s = 1/\tau$ é a frequência de amostragem. A função amostrada $\hat{x}(t)$ é formada pelo produto da função inicial $x(t)$, de espectro limitado, com a função $s(t)$. Pode-se portanto escrever

$$\hat{x}(t) = x(t)s(t) = \sum_{n=-\infty}^{\infty} x(n\tau)\delta(t - n\tau), \quad (2-1.23)$$

e o espectro desta função amostrada é evidentemente

$$\hat{X}(f) = X(f) * S(f) = \frac{1}{\tau} \sum_{k=-\infty}^{\infty} X(f - kf_s). \quad (2-1.24)$$

Pode-se ver desta maneira, que o espectro de $\hat{X}(f)$ se encontra a partir do espectro do sinal inicial, $X(f)$, retardando este de $\pm kf_s$, isto é, valores múltiplos da frequência de amostragem. Este processo de amostragem é ilustrado na figura 2.4.

Estes resultados foram obtidos considerando o caso particular em que a frequência de amostragem era suficientemente elevada em relação à frequência máxima do sinal $x(t)$, isto é, $f_s > W$. Observando a figura 2.4 torna-se evidente que, para que não exista sobreposição de dois espectros consecutivos, é necessário e suficiente que a frequência de amostragem f_s seja superior ou igual a $2W$, isto é, que

$$f_s > 2W. \quad (2-1.25)$$

Esta condição é absolutamente necessária para poder reconstituir o sinal $x(t)$ a partir de $\hat{x}(t)$ através da filtragem passa-baixo deste último. Neste caso

$$Y(f) = H_r(f)\hat{X}(f) = \hat{X}(f)\tau \text{rect}\left(\frac{f}{W}\right) = X(f), \quad (2-1.26)$$

quando $W \leq f_s/2$. Este processo de reconstituição está também representado na figura 2.4. Neste momento podemos estabelecer o teorema fundamental da amostragem, também chamado de Nyquist ou de Shanon:

Teorema

“Se uma função não contém componentes superiores a uma frequência W (hz), então essa função pode ser completamente determinada especificando o seu valor a uma série de pontos espaçados no máximo de $1/(2W)$ segundos.”

Evidentemente a transformada inversa de (2-1.26) é

$$x(t) = h_r(t) * \hat{x}(t) = h_r(t) * \sum_{n=-\infty}^{\infty} x(n\tau)\delta(t - n\tau), \quad (2-1.27)$$

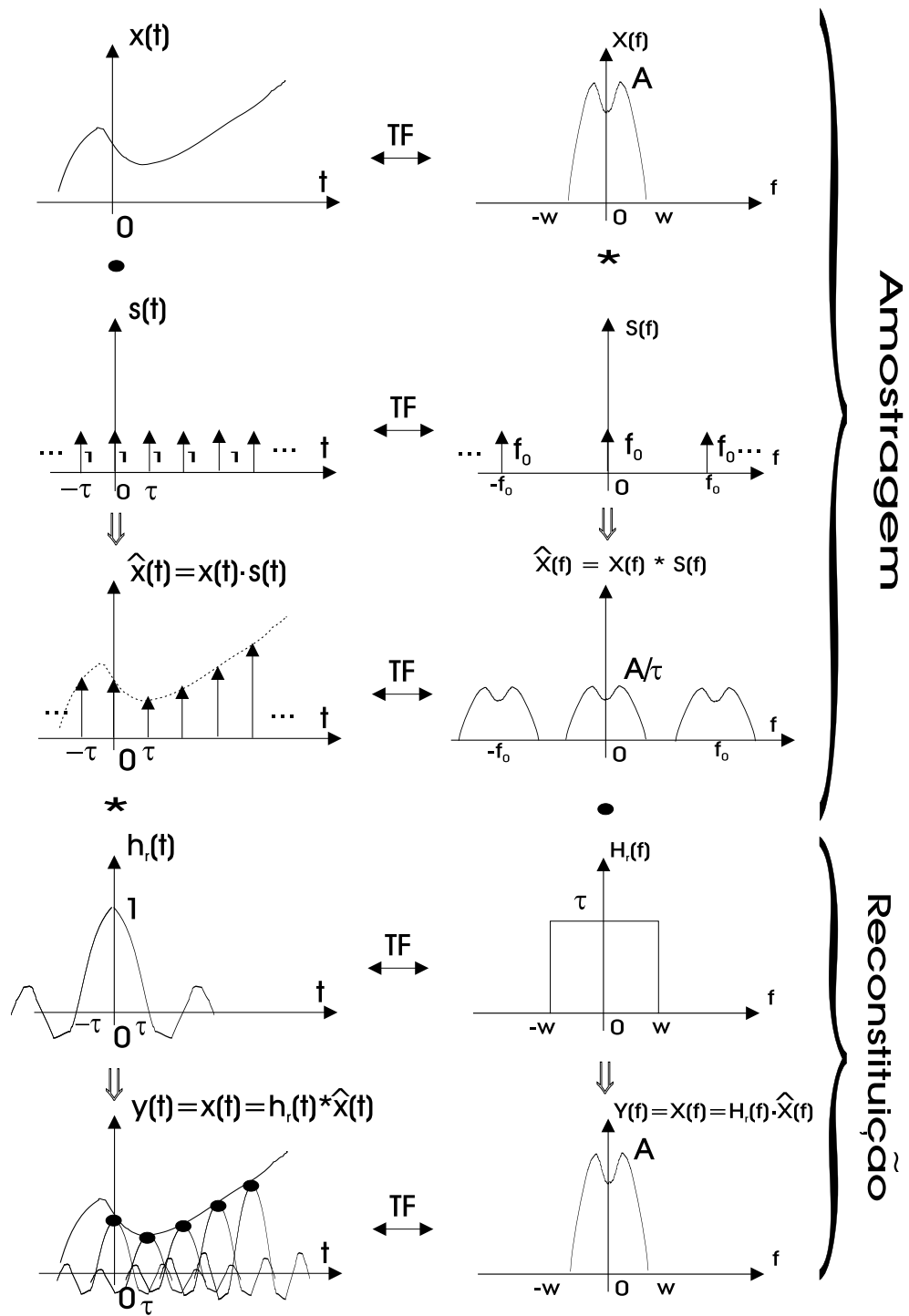


Figura 2.4: processo de amostragem e reconstituição.

que se pode também escrever

$$x(t) = \sum_{-\infty}^{\infty} x(n\tau)h_r(t - n\tau), \quad (2-1.28)$$

no caso do filtro passa-baixo ideal, a função de interpolação que é a TF^{-1} da função porta é

$$h_r(t - n\tau) = \tau \frac{\sin 2\pi W(t - n\tau)}{\pi(t - n\tau)}, \quad (2-1.29)$$

o que nos dá por substituição em (2-1.28) e com $\tau = 1/(2W)$,

$$x(t) = \sum_{-\infty}^{\infty} x(n\tau) \frac{\sin 2\pi W(t - n\tau)}{2\pi W(t - n\tau)}. \quad (2-1.30)$$

Esta formula permite-nos determinar a função contínua $x(t)$ a partir das amostras $x(n\tau)$.

Teorema de amostragem de sinais passa-banda

Já vimos o teorema de amostragem para as funções passa-baixo. No caso de uma função passa-banda, como por exemplo um sinal modulado, a frequência de amostragem não deve forçosamente ser \geq a duas vezes a frequência máxima do sinal. Isto deve-se ao facto que o espectro de um sinal passa-banda tem “buracos” com ganho nulo. Para estes sinais o teorema de amostragem escreve-se

Teorema

“Se uma função $x(t)$ tem um espectro passa-banda não nulo entre $f_c - W$ e f_c então a frequência de amostragem mínima é

$$f_s = \frac{2f_c}{m}, \quad (2-1.31)$$

onde m é o maior inteiro inferior a f_c/W . Atenção: todas as frequências superiores a f_s definida por (2-1.31) não são forçosamente utilizáveis devido ao problema de sobreposição de espectros.”

Podemos fazer a prova gráfica do teorema acima sem dificuldade. Consideremos o sinal $X(f)$ de banda W e frequência máxima f_c com $f_c/W = 2.5$ e escolhamos arbitrariamente $f_s = 1.1f_c$. O resultado pode ver-se na figura 2.5. Através de uma reconstrução gráfica deste tipo pode-se chegar aos vários valores de f_s que se podem obter para um dado f_c/W .

A figura 2.6 mostra um ábaco que permite determinar a frequência de amostragem mínima para sinais passa-banda a partir da frequência máxima f_c e da banda W do sinal.

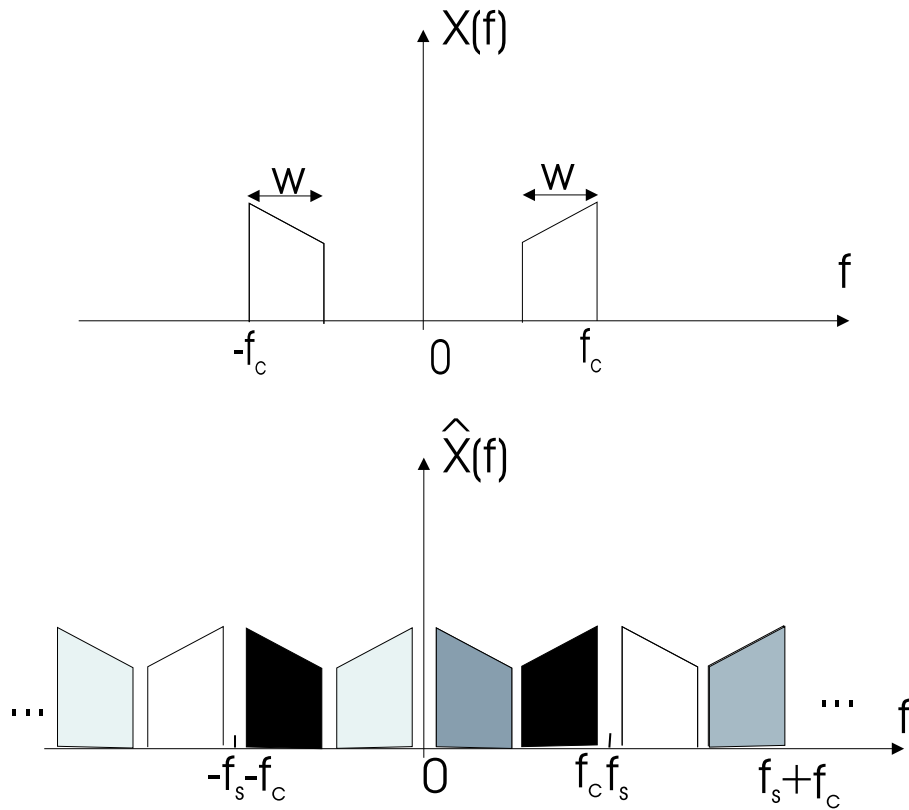


Figura 2.5: repetição espectral por amostragem de um sinal discreto.

Representação por modulação de impulsos

A partir do que foi dito acima, podemos deduzir que o sinal discreto ideal $x_n = x(n\tau)$ é

$$x_n = x(n\tau) = \int_{-\infty}^{+\infty} x(t)\delta(t - n\tau)dt \quad (2-1.32)$$

obtemos então o sinal discreto no tempo, representado pela sequência de amostras $\{x_n\}$. Inversamente o sinal amostrado $\hat{x}(t)$ pode ser representado por uma soma finita de *impulsos modulados em amplitude* pela sequência $\{x_n\}$,

$$\hat{x}(t) = \sum_{n=-\infty}^{\infty} x_n\delta(t - n\tau) \quad (2-1.33)$$

É esta última forma de representação que nos permite introduzir a noção de modulação de amplitude de impulsos (pulse amplitude modulation - PAM) da qual falaremos mais adiante.

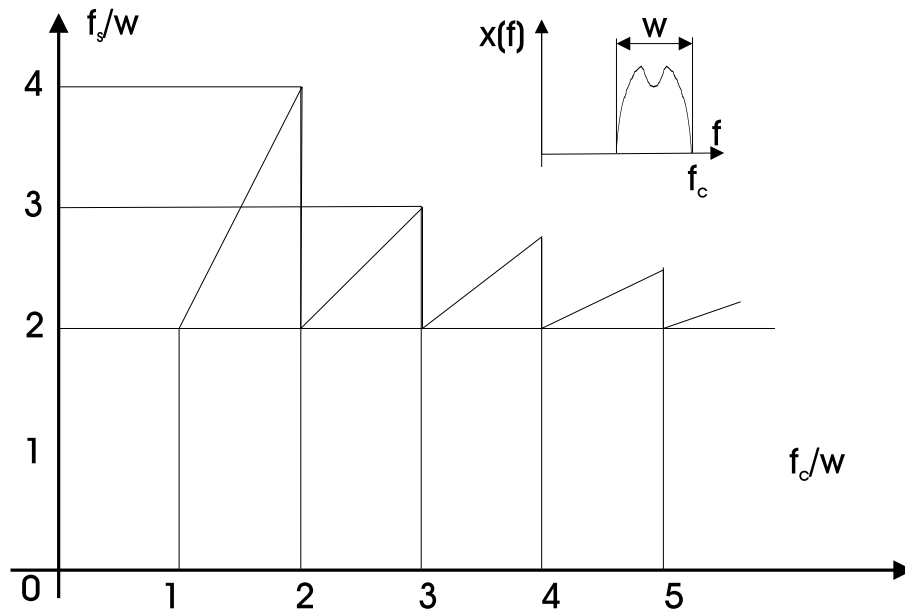


Figura 2.6: ábaco para determinar f_s no caso de sinais passa-banda.

2.1.4 Energia e potência

A energia de um sinal é uma quantidade fundamental nos sistemas de comunicação. Assim definimos a energia de um sinal por

$$E_x = \int_{-\infty}^{\infty} |x(t)|^2 dt, \quad \text{e} \quad E_x = \sum_{k=-\infty}^{\infty} |x_k|^2 \quad (2-1.34)$$

para sinais contínuos e discretos respectivamente. Sendo a potência igual à energia por unidade de tempo, só poderemos calcular a potência contida num sinal como sendo a energia média dada por

$$P_x = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} |x(t)|^2 dt, \quad \text{e} \quad P_x = \lim_{K \rightarrow \infty} \frac{1}{K} \sum_{k=-K/2}^{K/2-1} |x_k|^2. \quad (2-1.35)$$

2.1.5 Sinais passa-banda e a transformada de Hilbert

Como já tivemos a ocasião de mencionar na introdução, um dos papéis desempenhados pelos sistemas de comunicação colocados no emissor, é o de adaptar o sinal ao canal de comunicação. Frequentemente, uma das características essenciais do canal de comunicação é de que a sua banda de frequências é limitada em torno a um determinado valor: dizemos que se trata de um canal "passa-banda" - é o caso do canal rádio-frequência. O sistema de comunicação emissor terá então de

transformar o sinal contendo a mensagem a emitir num sinal passa-banda compatível com o canal de transmissão. Vamos agora descrever o formalismo matemático que permite representar sinais passa-banda e a sua manipulação ao longo da cadeia de transmissão.

Vamos considerar o sinal passa-banda $s(t)$ com o espectro $S(f)$. Primeiramente vamos considerar um novo sinal que contenha apenas a parte positiva do espectro de $s(t)$,

$$S_+(f) = 2u(f)S(f), \quad (2-1.36)$$

onde $u(f)$ é a função degrau unidade e o factor 2 tem em conta a conservação da energia, *i.e.*, $S_+(f)$ e $S(f)$ têm a mesma energia. No domínio do tempo temos

$$s_+(t) = \int_{-\infty}^{\infty} S_+(f)e^{j2\pi ft}df \quad (2-1.37)$$

$$= \text{TF}^{-1}[2u(f)] * \text{TF}^{-1}[S(f)]. \quad (2-1.38)$$

Visto que (ver apêndice A.1)

$$\text{TF}^{-1}[2u(f)] = \delta(t) + \frac{j}{\pi t} \quad (2-1.39)$$

temos

$$s_+(t) = \left[\delta(t) + j \frac{1}{\pi t} \right] * s(t). \quad (2-1.40)$$

O sinal $s_+(t)$ é chamado *pre-envelope* ou *sinal analítico* de $s(t)$, e não é mais do que uma versão complexa de $s(t)$, cuja parte real é o próprio $s(t)$ e a parte imaginária é a Transformada de Hilbert de $s(t)$, que defenimos de seguida. A partir do segundo termo do segundo membro de (2-1.40) define-se

$$\hat{s}(t) = \frac{1}{\pi t} * s(t) \quad (2-1.41)$$

$$= \frac{1}{\pi} \int \frac{s(\tau)}{t - \tau} d\tau \quad (2-1.42)$$

como sendo a *Transformada de Hilbert* de $s(t)$, cuja notação é

$$\text{H}[s(t)] = \frac{1}{\pi} \int \frac{s(\tau)}{t - \tau} d\tau. \quad (2-1.43)$$

Uma forma alternativa de ver a transformada de Hilbert é de considerar que o sinal $\hat{s}(t)$ pode ser visto como o sinal de saída de um filtro cuja resposta impulsiva é

$$h(t) = \frac{1}{\pi t}, \quad -\infty < t < \infty \quad (2-1.44)$$

excitado pelo sinal $s(t)$ à entrada. Este filtro é então chamado um transformador de Hilbert e a sua resposta em frequência escreve-se

$$H(f) = \text{TF}[h(t)] = -j \text{sgn}(f) = \begin{cases} -j & f > 0 \\ 0 & f = 0 \\ j & f < 0 \end{cases}. \quad (2-1.45)$$

Podemos notar que a função de transferência do filtro de Hilbert é tal que o seu módulo é $|H(f)| = 1$ e a sua fase é

$$\Phi(f) = \begin{cases} -\frac{\pi}{2} & f > 0 \\ \frac{\pi}{2} & f < 0 \end{cases} \quad (2-1.46)$$

o que nos faz dizer que este filtro é basicamente um desfasador puro de $\pi/2$ para todas as frequências do sinal de entrada, enquanto a amplitude não é alterada.

2.1.6 Passagem em banda-base

O sinal analítico $s_+(t)$ é um sinal passa-banda, *i.e.*, toma valores diferentes de zero para uma banda de frequências em torno a uma frequência f_c , e é zero para todos os outros valores de f . Para colocar o sinal $s_+(t)$ em banda base, *i.e.*, numa banda de frequências em torno a $f = 0$, teremos de fazer uma translação de f_c Hz. Assim o sinal em banda base equivalente a $s_+(t)$ será (no domínio da frequência) $S_{bb}(f) = S(f + f_c)$ ou seja no tempo

$$s_{bb}(t) = s_+(t)e^{-j2\pi f_c t} \quad (2-1.47)$$

$$= [s(t) + j\hat{s}(t)]e^{-j2\pi f_c t} \quad (2-1.48)$$

e de forma equivalente

$$[s(t) + j\hat{s}(t)] = s_{bb}(t)e^{j2\pi f_c t} \quad (2-1.49)$$

de onde fazendo $s_{bb}(t) = x(t) + jy(t)$ e igualando partes reais e imaginárias de (2-1.49) obtemos

$$s(t) = x(t) \cos(2\pi f_c t) - y(t) \sin(2\pi f_c t), \quad (2-1.50)$$

$$\hat{s}(t) = x(t) \sin(2\pi f_c t) + y(t) \cos(2\pi f_c t). \quad (2-1.51)$$

A equação (2-1.50) é uma representação do sinal passa-banda $s(t)$ no qual as suas componentes em banda-base $x(t)$ e $y(t)$ se encontram como moduladoras do cos e do sin às frequências centrais f_c ; $x(t)$ e $y(t)$ são chamadas as componentes em quadratura do sinal $s(t)$ visto que as funções que elas modulam estão desfasadas de $\pi/2$. Alternativamente, utilizando a representação exponencial dum número complexo, podemos escrever

$$s_l(t) = a(t)e^{j\theta(t)} \quad (2-1.52)$$

onde

$$a(t) = \sqrt{x^2(t) + y^2(t)} \quad (2-1.53)$$

$$\theta(t) = \tan^{-1} \frac{y(t)}{x(t)} \quad (2-1.54)$$

onde $a(t)$ e $\theta(t)$ são chamados o envelope e a fase de $s(t)$, respectivamente.

2.2 Sinais aleatórios

Em muitos problemas físicos é necessário associar um valor numérico ao resultado de uma experiência aleatória. Por exemplo, o sinal da figura 2.7 poderia representar a voltagem medida aos bornos de uma resistência geradora de ruído. Poderia também representar a temperatura, em graus centígrados, medida num dado local em função do tempo. Em ambos os casos ao resultado de uma experiência imprevisível foi associado um valor numérico. Este conceito, extremamente importante na modelização de sinais físicos não determinísticos, será introduzido neste capítulo. Aconselha-se aqui a leitura dos conceitos de base sobre probabilidades elementares, variáveis aleatórias e a sua caracterização assim como relembrar as leis de distribuição mais usuais (apêndice C).

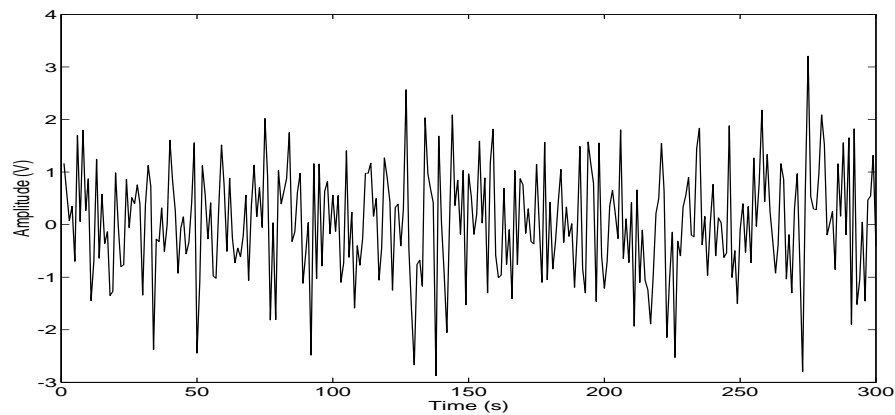


Figura 2.7: sinal aleatório.

2.2.1 Noção de processo aleatório

A noção de processo aleatório está na base da representação matemática de sinais aleatórios. Um sinal temporal aleatório, que nós representaremos $X(t, w)$ será, a

cada instante t , uma VA cuja distribuição dependerá do acontecimento w . Assim o processo aleatório $X(t, w)$ será:

- para t fixo, uma VA dependente do acontecimento w .
- para w fixo, uma função contínua de variável contínua t , chamada usualmente trajectória.

Deste modo $X(t, w)$ pode ser assimilado a uma “família” de trajectórias dependentes da realização de w . Se t é uma variável contínua, geralmente definida em $[-\infty, \infty]$, diz-se que $X(t, w)$ é um processo estocástico contínuo. Se t toma valores num conjunto finito de pontos t_i , diz-se que $X(t_i, w)$ é um processo estocástico discreto.

Considerando um processo aleatório discreto $X(t_i, w)$ que toma valores para os instantes $t_i; i = 1, \dots, n$ poderemos formar um vector aleatório de dimensão n , $\mathbf{X}^t = [X(t_1), X(t_2), \dots, X(t_n)]$. Apesar de ainda não termos estudado conjuntos de variáveis aleatórias, ou variáveis aleatórias multidimensionais, podemos desde já adiantar que o vector aleatório \mathbf{X} será completamente definido em termos estatísticos, por uma função de distribuição $F(\mathbf{x})$ que será função de $x_1 = X(t_1)$, $x_2 = X(t_2)$, ..., $x_n = X(t_n)$. Óbviamente, o objectivo aqui é o de substituir o estudo da família infinita de trajectórias no intervalo $t_1 \leq t \leq t_n$ por um número finito de VA's. Esta substituição não é matematicamente correcta mas justifica-se através das condições de amostragem temporal (que já estudamos) por um lado, e através de condições de “estacionaridade” em termos estatísticos (que veremos mais à frente) por outro. Isto é, se o sinal aleatório for amostrado obedecendo às regras de amostragem mínimas já enunciadas, e se a sua função de distribuição $F(\mathbf{x})$ tiver certas propriedades no intervalo considerado $[t_1, t_n]$, então a redução de $X(t, w)$ a \mathbf{X} é justificada. Uma das condições estatísticas de base é que

$$F(x_i) = F(x), \quad \forall i \in [1, n] \quad (2-2.1)$$

ou seja que a função de distribuição de x seja independente de t . Na sequência deste capítulo, para simplificar, utilizaremos a notação $X(t)$ em vez de $X(t, \omega)$, como fizemos no caso das VA's.

2.2.2 Média e variância de um processo aleatório

A média de um sinal aleatório é definida por

$$m(t) = E[X(t)], \quad (2-2.2)$$

onde se pode observar que neste caso a média $m(t)$ é uma função do tempo t . Se a função de distribuição de X for independente do tempo, então $m(t) = m$, será

também independente do tempo. A variância de um sinal real é dada por

$$V[X(t)] = \sigma^2(t) = E[\{X(t) - m(t)\}^2] = E[X^2(t)] - m^2(t). \quad (2-2.3)$$

A média e a variância de sinais aleatórios têm um importante significado físico. Na maior parte dos casos teremos que $m(t) = 0$ e nos casos em que $m(t) \neq 0$ então poderemos achar conveniente considerar o sinal aleatório $X(t) - m(t)$. A variância do sinal aleatório de média nula será então $\sigma^2(t) = E[X^2(t)]$, ou seja, representará a sua potência média.

2.2.3 Covariância e correlação

Consideremos agora dois instantes t_1 e t_2 . As duas variáveis aleatórias $X_1 = X(t_1)$ e $X_2 = X(t_2)$ podem ser comparadas em termos estatísticos através da função de covariância que se define por

$$\gamma(t_1, t_2) = COV[X_1, X_2] = E[(X_1 - m_1)(X_2 - m_2)] = E[X_1 X_2] - m_1 m_2, \quad (2-2.4)$$

onde $m_1 = m(t_1)$ e $m_2 = m(t_2)$. Óbviamente a variância σ^2 é obtida quando $t_1 = t_2 = t$, i.e., através de

$$\sigma^2(t) = \gamma(t, t) = COV[X, X] \quad (2-2.5)$$

Em muitos casos, e em particular em caso de *estacionaridade*, teremos que a covariância será uma função apenas do intervalo $\tau = t_1 - t_2$ e então

$$\gamma(t_1, t_2) = \gamma(\tau), \quad (2-2.6)$$

à qual se chama geralmente função de correlação do sinal aleatório $X(t)$. Por outras palavras a função de correlação de um sinal aleatório (estacionário) é definida por

$$\gamma(\tau) = \gamma(t, t - \tau) = E[\{X(t) - m\}\{X(t - \tau) - m\}]. \quad (2-2.7)$$

Definiremos mais adiante o conceito de estacionaridade.

Pela primeira vez falaremos aqui de variáveis aleatórias complexas, numa preocupação de maior generalidade dos resultados e porque são necessárias em muitas aplicações. Uma VA complexa Z não é mais do que um conjunto de duas VA's reais ligadas por $Z = X + jY$. Todas as propriedades básicas podem ser deduzidas das VA's reais tendo em conta que $Z^2 = |Z|^2$. A extensão é evidente para sinais aleatórios complexos a partir de sinais aleatórios reais. No caso da função de correlação, um sinal aleatório estacionário complexo de valor médio m admite por função de correlação

$$\gamma(\tau) = E[X(t)X^*(t - \tau)] - |m|^2, \quad (2-2.8)$$

que no caso dos sinais reais (2-2.8) se reduz a

$$\gamma(\tau) = E[X(t)X(t - \tau)] - m^2. \quad (2-2.9)$$

Como consequência da estacionaridade é claro que podemos substituir t por $t + \tau$ nas relações (2-2.8) e (2-2.9) e notar que a função de correlação de um sinal real é uma função par de τ e que em especial a variância $\sigma^2 = \gamma(0)$.

No caso das VA's discretas podemos fazer a extensão óbvia da função de covariância

$$\gamma(k_1, k_2) = E[X(k_1)X^*(k_2)] - m(k_1)m^*(k_2), \quad (2-2.10)$$

e para a função de correlação

$$\gamma(p) = E[X(k)X^*(k - p)] - |m|^2, \quad (2-2.11)$$

onde p, k_1 e k_2 são números inteiros.

A função de covariância dá uma medida física do carácter aleatório de um sinal. Isto é mais claro no caso da função de correlação: se ela é nula significa que os sinais $X(t)$ e $X(t - \tau)$ são não correlados qualquer que seja o valor de t , veremos mais tarde que isto significa que o sinal $X(t)$ perdeu a memória do seu valor a $t - \tau$.

Algumas outras propriedades da função de covariância são:

(a) Existência: a partir da desigualdade de Schwartz¹ temos que

$$|\gamma(t_1, t_2)|^2 \leq \gamma(t_1, t_1)\gamma(t_2, t_2), \quad (2-2.12)$$

e portanto a existência da função de covariância está garantida sempre que $\gamma(t, t) < \infty$.

(b) Simetria: para sinais reais podemos dizer que $\gamma(t_1, t_2) = \gamma(t_2, t_1)$. No caso estacionário, em que a função de correlação é apenas uma função de $\tau = t_2 - t_1$, temos que $\gamma(\tau) = \gamma^*(-\tau)$ se os sinais forem complexos e $\gamma(\tau) = \gamma(-\tau)$ se os sinais forem reais.

(c) Valores extremos: no caso estacionário temos que

$$|\gamma(\tau)| \leq \gamma(0), \quad (2-2.13)$$

por outro lado quando $\tau \rightarrow \infty$ os sinais tendem a ser descorrelacionados e por isso podemos dizer que em geral

$$\lim_{\tau \rightarrow \infty} \gamma(\tau) = 0. \quad (2-2.14)$$

¹ $E[AB]^2 \leq E[A^2]E[B^2]$

- (d) Adição de covariâncias: sejam dois sinais $X_1(t)$ e $X_2(t)$ não correlados e de covariâncias respectivas γ_1 e γ_2 . A covariância do sinal $X(t) = X_1(t) + X_2(t)$ no intervalo (t_i, t_j) é

$$\gamma(t_i, t_j) = \gamma_1(t_i, t_j) + \gamma_2(t_i, t_j). \quad (2-2.15)$$

- (e) Produto de covariâncias: consideremos dois sinais aleatórios de média nula e independentes. Devido à propriedade de independência o sinal $Y(t) = X_1(t)X_2(t)$ tem como covariância

$$\gamma_y(t_i, t_j) = \gamma_1(t_i, t_j)\gamma_2(t_i, t_j). \quad (2-2.16)$$

- (f) Tempo de correlação: como foi dito atrás para sinais reais a função de correlação tende para zero quando o tempo tende para infinito. Esta variação pode ser mais ou menos rápida e é costume caracterizá-la por um parâmetro chamado tempo de correlação, τ_c , que pode ser definido de várias formas. Na prática para $t > \tau_c$ os sinais são ditos não correlados.

2.2.4 Estacionaridade e ergodicidade

O conceito de estacionaridade foi referido mais do que uma vez neste capítulo sem que uma definição precisa tenha sido dada. O conceito mais importante de estacionaridade prende-se com a noção de invariância por translação temporal. No caso dos sinais aleatórios diz-se que o sinal aleatório $X(t)$ é estacionário no sentido estrito se as suas propriedades estatísticas são invariáveis por qualquer translação da origem do tempo. Invariância temporal no sentido estrito requer que a função de distribuição seja idêntica a qualquer instante de $-\infty$ a $+\infty$. Esta propriedade é dificilmente verificável na prática e uma das maneiras para se poder aliviar um pouco esta definição é de considerar que para muitos problemas é suficiente que apenas os momentos até à segunda ordem sejam temporalmente invariantes. Diz-se nesse caso que o sinal é estacionário de segunda ordem ou estacionário no sentido lato. O conceito de estacionaridade reveste-se de uma importância primordial em análise de sinais sendo no entanto um conceito puramente teórico devido ao facto de nunca poder ser verificado na prática. Com efeito, é impossível verificar a estacionaridade na prática porque toda a experiência no mundo real tem forçosamente uma duração limitada. Noções de semi-estacionaridade e estacionaridade local para sinais limitados no tempo são frequentes e geralmente aceites na prática.

A noção de ergodicidade é bastante mais subtil. Começemos por considerar um processo aleatório $X(t, w)$. Para $w = w_0$ $X(t, w_0)$ é uma simples função do tempo à qual chamamos trajectória. Na prática é essa trajectória que é medida ao longo do tempo. É comum calcular-se uma média temporal no intervalo de observação T a partir de

$$m(T; \omega_0) = \frac{1}{T} \int_0^T X(t, \omega_0) dt, \quad (2-2.17)$$

que não deve ser confundida com a esperança matemática do sinal $X(t)$. A questão é de saber em que medida é que a média temporal (2-2.17) está relacionada com a média de conjunto ou esperança matemática $E[X(t)]$? A noção de ergodicidade prende-se exactamente com a resposta a esta questão, postulando que para um sinal *ergódico*, existe equivalência, para largos valores de T , entre as médias de conjunto e as médias temporais. A ergodicidade é uma propriedade importante na prática e que, mais uma vez é difícil (senão impossível) de verificar devido à impossibilidade de efectuar uma série infinita de tiragens aleatórias de modo a poder calcular uma média de conjunto estatisticamente válida. É de notar que ergodicidade não implica necessariamente estacionaridade e vice-versa, apesar de se encontrarem quase sempre associadas na prática.

2.2.5 Energia e potência

Como no caso dos sinais determinísticos, qualquer quantidade que faça intervir uma relação quadrática do sinal contém uma noção de energia. Para um sinal estocástico complexo e de média nula $X(t, \omega)$, a grandeza

$$P(t, \omega) = |X(t, \omega)|^2, \quad (2-2.18)$$

é a potência aleatória instantânea. Os sinais aleatórios para os quais $P(t, \omega) < \infty$ são chamados de segunda ordem e, de acordo com (2-2.9), $P(t) = E[P(t, \omega)] = \gamma(t, t)$. Mais, se além disso o sinal $X(t, \omega)$ for de segunda ordem e estacionário então a média de potência instantânea torna-se independente do tempo e

$$P(t) = P = \gamma(0), \quad (2-2.19)$$

que não é mais do que a variância do sinal em questão. Por outras palavras, para sinais de média nula a potência média instantânea é igual à variância. O conceito de energia é um pouco mais subtil de tal modo que, sabendo que a energia é a potência durante um determinado intervalo de tempo, a energia total contida num sinal é

$$J(\omega) = \int |X(t, \omega)|^2 dt, \quad (2-2.20)$$

onde o integral vai de $-\infty$ a ∞ . Esta energia não será geralmente finita. Porém, num intervalo de tempo finito ela será também finita. A energia média (em termos estatísticos) é

$$J = E[J(\omega)] = \int P(t) dt = \int \gamma(t, t) dt. \quad (2-2.21)$$

Sinais estacionários, para os quais $P(t) = P = \text{constante}$, a energia J é infinita. É idêntico ao caso observado para os sinais determinísticos periódicos para os quais a energia também é infinita. Neste ponto é útil introduzir a noção de potencia média, calculada num intervalo de tempo limitado, que será

$$P_T(\omega) = \frac{1}{T} \int_0^T P(t, \omega) dt, \quad (2-2.22)$$

a partir da qual podemos, sob condição de ergodicidade, dizer que

$$P(t) = \lim_{T \rightarrow \infty} P_T(\omega). \quad (2-2.23)$$

É óbvio que, se a energia (2-2.21) do sinal for finita, então a potência média $P_T(\omega)$ tenderá para zero quando $T \rightarrow \infty$, que é a mesma noção que já vimos para sinais determinísticos, para os quais se tiverem uma energia finita a sua potência média temporal será nula.

2.2.6 Filtragem linear de sinais aleatórios

Já vimos que um sistema (ou filtro) linear é um sistema para o qual a relação entre a entrada e a saída se escreve através de uma equação de convolução. Isto, no quadro de sinais determinísticos. Vejamos agora no caso de sinais aleatórios. Nesse caso podemos escrever

$$y(k) = \sum h(l)x(k-l), \quad (2-2.24)$$

para o caso de sinais discretos, onde $x(k)$ é o sinal de entrada, $h(k)$ é a resposta impulsiva do filtro e $y(k)$ é o sinal de saída. Poderíamos ter escrito também (2-2.24) para o caso de sinais contínuos sem alteração do resultado que se pretende demonstrar neste capítulo. Para podermos admitir (2-2.24) teremos que considerar que o filtro é absolutamente estável e que o sinal de entrada é de média finita de modo a assegurar que o somatório (geralmente calculado de $-\infty$ a ∞) seja finito. O problema interessante aqui é de, conhecendo as propriedades estatísticas de $x(k)$, determinar as do sinal de saída $y(k)$ que será, neste caso, também aleatório. Como vimos anteriormente uma VA encontra-se completamente caracterizada estatisticamente pela sua função de distribuição. No caso de um sinal aleatório, como $y(k)$, a sua caracterização estatística faz-se pela família de funções de distribuição das VA's de que é formado, no intervalo de tempo considerado. Determinar a família de funções de distribuição do sinal de saída a partir das distribuições do sinal de entrada é uma tarefa extremamente difícil, senão impossível, na maioria dos casos. Por exemplo, no caso em que filtro é FIR, o sinal de saída é uma soma finita de termos aleatórios, e portanto uma soma finita de VA's. Utilizando regras simples de combinação de VA's poderemos calcular a função característica de um VA Y tal que

$$Y = \sum_{i=1}^N X_i. \quad (2-2.25)$$

Agora, passar da função característica à função de distribuição é uma operação que requer uma transformada de Fourier inversa multidimensional que não é de modo nenhum fácil de obter em geral. Se as funções de distribuição são, em geral, extremamente difíceis de obter, as esperanças matemáticas são, elas, relativamente fáceis. Assim,

$$m_y(k) = \sum h(l)m_x(k-l), \quad (2-2.26)$$

ou seja, se a entrada é de média nula a saída também será de média nula. No caso da segunda ordem,

$$\gamma_y(k_1, k_2) = \sum_{l_1} \sum_{l_2} h(l_1)h^*(l_2)\gamma_x(k_1 - l_1, k_2 - l_2). \quad (2-2.27)$$

Finalmente no caso de sinais estacionários, o que para os filtros significa que são invariantes temporalmente, podemos escrever

$$m_y = m_x \sum h(k), \quad (2-2.28)$$

e que

$$\gamma_y(k) = \sum_{l_1} \sum_{l_2} h(l_1)h(l_2)^*\gamma_x(k + l_2 - l_1), \quad (2-2.29)$$

o que tem a óbvia vantagem de podermos dizer que também a média e a covariância do sinal de saída são estacionárias.

2.2.7 Ruído branco

Apesar de ser um pouco cedo neste curso para introduzir a noção de ruído na sua generalidade torna-se importante para a compreensão dos próximos capítulos a introdução da noção de ruído branco que se justifica pela sua importância na prática.

Ruído branco é por si só um nome extremamente utilizado na literatura mas que é, devido a um abuso de linguagem, um pouco enganador. Deveria dizer-se *sinial branco*. Na realidade trata-se de um modelo estatístico que descreve um certo tipo de sinais que têm determinadas propriedades. Quase sempre essas propriedades são indesejáveis num sinal e por isso chama-se-lhe *ruído* daí que quase todos os *siniais brancos* são *ruídos brancos* o que justifica o nome. Um sinal aleatório estacionário de média nula, $x(k)$, é *branco* se a sua função de correlação se escreve

$$\gamma(k) = \sigma^2\delta(k), \quad (2-2.30)$$

onde σ^2 é a potência média (ou variância) do sinal. A expressão (2-2.30) significa que $x(k)$ e $x(l)$ são descorrelados para $l \neq k$ o que, por outras palavras, quer dizer que o sinal $x(k)$ é uma sequência de VA's não correladas e com a mesma variância. Veremos no próximo capítulo que a partir de (2-2.30) podemos calcular a densidade espectral de potência (DEP ou PSD=power spectral density) de $x(k)$ através da DFT de $\gamma(k)$ obtendo assim $P_x(f) = P = \sigma^2$. O nome de *branco* vem da óptica onde a cor branca tem um espectro muito largo e constante, por oposição à luz colorida, que tem um espectro mais estreito. Existe uma diferença entre não correlação e independência¹,

¹Lembramos que independência entre duas variáveis aleatórias significa que o conhecimento de uma de entre elas não nos traz nenhuma informação sobre a outra, i.e., $Pr(A/B) = Pr(A)$. Não correlação significa desigualdade ou perda de memória entre duas VA's.

da qual já falámos, e que nos leva naturalmente à noção de ruído branco no sentido estrito, que consiste num ruído branco como definido acima, mas onde todas as VA's seguem a mesma lei de probabilidade. Este é evidentemente o mais simples sinal aleatório que possa existir. Em particular, e veremos a sua importância mais tarde, se a lei de probabilidade for gaussiana dizemos estar em presença de ruído branco gaussiano. Para exemplificar de um modo gráfico estes conceitos representámos na

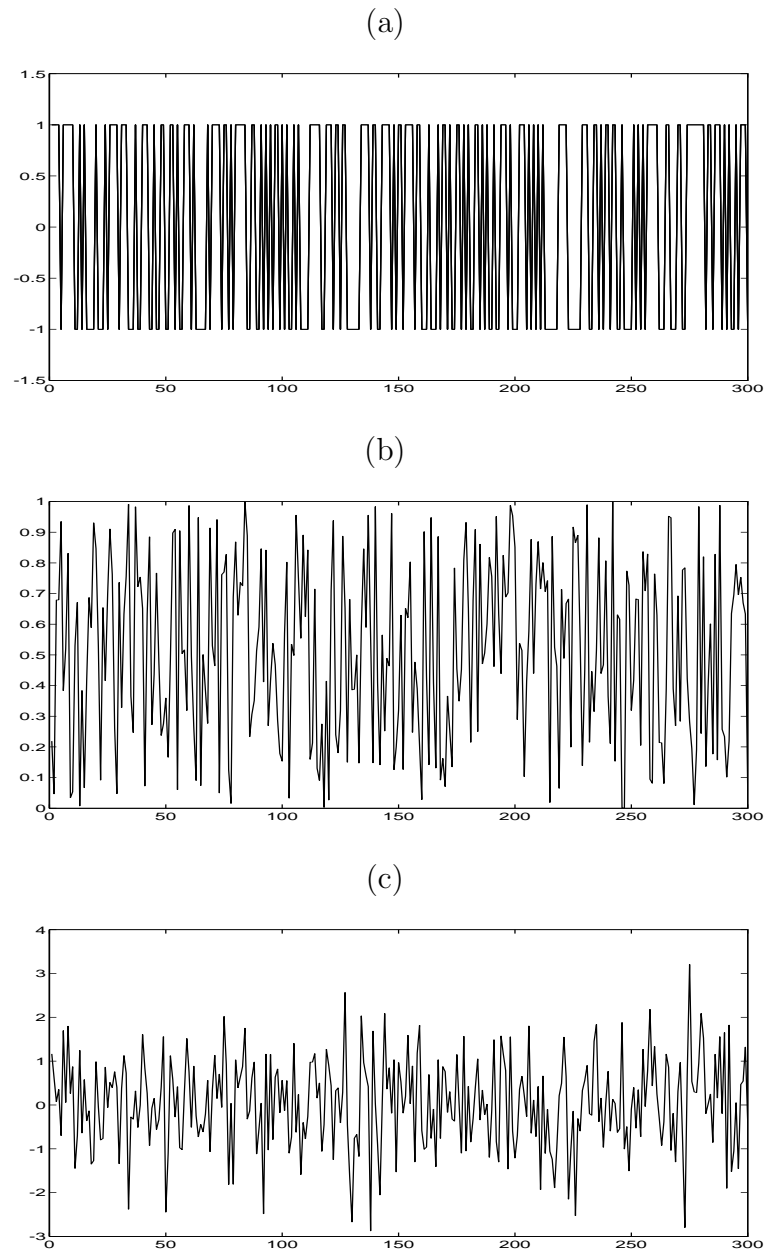


Figura 2.8: ruído branco: binário (a), uniforme (b) e gaussiano (c).

figura 2.8 algumas sequências de ruído branco: em (a) com amplitude discreta, em (b) distribuído segundo uma lei uniforme e em (c) segundo uma lei gaussiana. Os sinais da figura 2.8 foram gerados com base num gerador de números aleatórios existente em qualquer computador. Trata-se na realidade, como é de conhecimento geral, de uma sequência pseudo-aleatória pois esta repete-se ao fim de um certo número de tiragens, que no entanto é suficientemente elevado, para podermos dizer que são quasi-independentes. A questão inversa que se pode fazer quando observamos uma sequência aleatória é se se trata de um sinal branco ou não? Trata-se de um problema de interesse considerável na prática. Pode-se, por exemplo, calcular a sua função de correlação através da média temporal (o que implica uma hipótese de ergodicidade) e determinar se obedece a (2-2.30).

Não existe nenhum método simples para verificar se se trata de uma sequência de VA's independentes. Existem testes estatísticos (que não descreveremos) que nos permitem afirmar com algum grau de confiança se se trata ou não de uma sequência de VA's independentes. Estudaremos mais à frente a filtragem linear de ruídos brancos e deduziremos as condições de transformação de um sinal não branco, num sinal branco através de filtragem que se designa pelo termo anglo-saxónico “whitening filter”. Não falaremos no entanto de ruído branco contínuo pois tem bastante menos interesse prático e, apesar de ter um tratamento semelhante, requer alguns conhecimentos sobre cálculo integral que estão fora do âmbito desta disciplina.

2.2.8 Introdução à estimação espectral

As noções de estacionaridade e ergodicidade introduzidas no capítulo anterior são de importância fundamental na estimação do conteúdo espectral dos sinais aleatórios. Intuitivamente, podemos compreender que, em princípio, só fará sentido falar de componentes espectrais de um sinal que tem um determinado grau de repetitividade temporal. Esta noção intuitiva de repetitividade só existe em sinais estacionários e é por essa razão que não faz sentido calcular o espectro de sinais não estacionários.

A primeira ideia que é preciso ter bem presente é de que, para um sinal aleatório $x(t)$, não limitado no tempo a sua transformada de Fourier não existe, pois o integral não converge entre $[-\infty, \infty]$. Veremos mais à frente como definir a densidade espectral de um sinal aleatório quando este é limitado no tempo. Por enquanto, podemos definir a densidade espectral de potência de um sinal aleatório **estacionário** $x(t)$ através de

$$P_{xx}(f) = \int_{-\infty}^{\infty} r_{xx}(\tau) e^{-j2\pi f\tau} d\tau \quad (2-2.31)$$

onde $r_{xx}(\tau)$ é a função de autocorrelação do sinal $x(t)$ definida por

$$r_{xx}(\tau) = E[x(t)x(t - \tau)]. \quad (2-2.32)$$

A relação (2-2.31) é fundamental em teoria do sinal e toma o nome de teorema

da autorrelação ou teorema de Wiener-Khinchine (WK). A quantidade $P_{xx}(f)$ é chamada espectro de potência ou densidade espectral de potência, que é um termo mais exacto pois a sua unidade exprime-se em watts/Hertz. Por isso a potência total (em watts) do sinal $x(t)$, por exemplo na banda $[f_1, f_2]$ escreve-se

$$P = \int_{f_1}^{f_2} P_{xx}(f) df. \quad (2-2.33)$$

Inversamente podemos calcular a função de autocorrelação de $x(t)$ a partir de $P_{xx}(f)$ através de

$$r_{xx}(t) = \int_{-\infty}^{\infty} P_{xx}(f) e^{j2\pi ft} df. \quad (2-2.34)$$

Visto que a função de autocorrelação e a densidade espectral de potência formam um par de TF, as suas propriedades são as mesmas da TF. Aconselha-se o leitor a rever estas propriedades na matéria de Sistemas e Sinais. A aplicação prática do teorema de WK encerra visivelmente um problema: *é que a função de autocorrelação (2-2.32) não pode ser calculada de forma exacta pois não dispomos de um número infinito de tiragens aleatórias do processo $x(t)$ para calcular a esperança matemática $E[.]$.* É aqui que entra a hipótese de ergodicidade, permitindo admitir que a média de conjunto pode ser substituída pela média temporal, *i.e.*, que para um processo ergódico temos que

$$E[x(t)x(t-\tau)] = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)x(t-\tau) dt. \quad (2-2.35)$$

Porém, o problema do cálculo da densidade espectral através do teorema de WK ainda não está completamente resolvido, pois na prática nós não dispomos de um intervalo temporal infinito e por isso (2-2.35) não nos é útil. Utilizando (2-2.35) num intervalo T finito obtemos, não a função de autocorrelação, mas uma sua estimada $\hat{r}_{xx}(t)$. É aqui que se apreende a razão fundamental pela qual o cálculo do espectro de sinais aleatórios reais não se chama simplesmente *cálculo espectral* mas sim *estimação espectral*. Para sinais reais nós nunca teremos acesso à verdadeira densidade espectral mas apenas a uma estimativa dessa densidade espectral.

Voltemos então ao problema da indefinição da TF de um sinal aleatório como impedimento para o cálculo de sua densidade espectral. Já vimos, quando da definição da TF, que a densidade espectral de energia se escreve, para o caso dos sinais determinísticos, como

$$S_{xx}(f) = |X(f)|^2. \quad (2-2.36)$$

O que impede o seu uso no caso dos sinais aleatórios é o facto de $X(f)$ não existir para o caso dos sinais aleatórios definidos em $[-\infty, \infty]$. Porém, acabámos de ver que os sinais reais são forçosamente observados num intervalo de tempo limitado e que portanto tem uma energia finita. Consideremos um intervalo de observação $[0, T]$. A energia total de $x(t)$ neste intervalo é dada por

$$\int_0^T x^2(t) dt = \int_{-\infty}^{\infty} x^2(t) dt = \int_{-\infty}^{\infty} |X(f)|^2 df. \quad (2-2.37)$$

Esta última igualdade vem do teorema de Parseval que postula que a energia total é conservada quando da passagem do domínio do tempo para o domínio da frequência. A potência total no mesmo intervalo escreve-se então, tendo em conta (2-2.37),

$$P_T = \frac{1}{T} \int_{-\infty}^{\infty} |X(f)|^2 df, \quad (2-2.38)$$

que também se pode escrever fazendo $T \rightarrow \infty$ e invertendo a ordem do limite e do integral

$$P = \int_{-\infty}^{\infty} \left[\lim_{T \rightarrow \infty} \frac{|X(f)|^2}{T} \right] df. \quad (2-2.39)$$

Comparando esta última equação com (2-2.33) (com bornos ∞) podemos, por identificação tirar que a densidade espectral de potência se escreve

$$P_{xx}(f) = \lim_{T \rightarrow \infty} \frac{1}{T} \left| \int_0^T x(t) e^{-j2\pi ft} dt \right|^2. \quad (2-2.40)$$

O termo

$$S_T(f) = \frac{1}{T} \left| \int_0^T x(t) e^{-j2\pi ft} dt \right|^2, \quad (2-2.41)$$

representa a potência média no intervalo T . Para que

$$P_{xx}(f) = \lim_{T \rightarrow \infty} S_T(f), \quad (2-2.42)$$

seja válido é necessário ainda que $S_T(f)$ seja um estimador consistente, *i.e.*, que quando $T \rightarrow \infty$

$$E[S_T(f)] = P_{xx}(f), \quad \text{e} \quad \sigma_{S_T}^2 = 0 \quad (2-2.43)$$

uma variância nula quer dizer que $S_T(f)$ toma um único valor (quando $T \rightarrow \infty$) que é $P_{xx}(f)$. Assim podemos escrever que

$$S_T(f) = E[S_T(f)], \quad \text{quando} \quad T \rightarrow \infty \quad \text{se} \quad \sigma_{S_T}^2 = 0 \quad (2-2.44)$$

e daí escrever a relação

$$P_{xx}(f) = \lim_{T \rightarrow \infty} \frac{1}{T} E[| \int_0^T x(t) e^{-j2\pi ft} dt |^2]. \quad (2-2.45)$$

As relações (2-2.31) e (2-2.45) constituem duas formas equivalentes (mas diferentes) de estimar a densidade espectral de potência de um processo aleatório. Estas duas formas enquadram-se naquilo que se chama a estimação espectral clássica por oposição a outros tipos de métodos ditos baseados em modelos paramétricos e/ou de alta resolução. A miríade de métodos de estimação espectral existentes hoje em dia é fruto do conceito fundamental que esta representa no processamento da informação e do esforço a ela dedicado pela comunidade científica durante os últimos 20/30 anos encontram-se descritos em inúmeros livros da especialidade e motivaram uma cadeira de opção deste curso (ver sebenta de Estimação Espectral e Aplicações em ftp.ualg.pt/users/sjesus/aulas/eea00.pdf).

2.2.9 Sinais cicloestacionários

Sinais cicloestacionários são sinais aleatórios cujas funções de autocorrelação são periódicas. Assim, a função de autocorrelação

$$r_{xx}(t + \tau, t) = E[x(t + \tau)X^*(t)], \quad (2-2.46)$$

é tal que

$$r_{xx}(t + \tau + kT, t + kT) = r_{xx}(t + \tau, t), \quad (2-2.47)$$

e portanto de período T . A caracterização destes sinais através da sua densidade espectral faz-se introduzindo a média temporal da sua função de autocorrelação num período que é

$$\bar{r}_{xx}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} r_{xx}(t + \tau, t) dt, \quad (2-2.48)$$

permitindo o cálculo da densidade espectral de potência média de um processo cicloestacionário definida por

$$P_{xx}(f) = \int_{-\infty}^{\infty} \bar{r}_{xx}(\tau) e^{-j2\pi f\tau} d\tau. \quad (2-2.49)$$

3 Canais de transmissão

3.1 Linhas de transmissão

Talvez o meio mais antigo e também mais utilizado até hoje para transmitir informação tenha sido o cabo bifilar. É no entanto cada vez menos utilizado hoje em dia devido à sua fraca capacidade para suportar grandes quantidades de informação. A sua utilização encontra-se quase praticamente restrita a curtas distâncias e a baixo débito. As razões principais devem-se essencialmente a: grande sensibilidade a interferências electromagnéticas, *cross-talk* e atenuação elevada em função da distância percorrida pelo sinal. O *cross-talk* é definido como sendo a interferência gerada num cabo por um outro na sua proximidade devido ao campo electromagnético gerado pela corrente que o atravessa. A atenuação é uma função da resistência própria do condutor e directamente ligado ao material empregue e à sua secção. Outros efeitos como o efeito de bobine e de condensador tornam a atenuação dependente da frequência, o que faz o cabo actuar como um filtro e limitar fortemente a banda de frequências do sinal que o pode atravessar e por isso a quantidade de informação.

O cabo coaxial é utilizado para transmissão de maiores quantidades de informação. A atenuação aumenta aproximadamente com a raiz da frequência do sinal transmitido e por isso requer alguma adaptação para altas frequências a longas distâncias. Uma vantagem típica dos cabos coaxiais é a sua grande imunidade a interferências electromagnéticas. Esta deve-se à construção concêntrica do cabo na qual o condutor exterior se encontra à massa e portanto faz o papel de gaiola de Faraday não deixando sair para o exterior quase nenhuma radiação. Encontram-se cabos coaxiais capazes de suportar uma banda de 60 MHz correspondente a cerca de 140 Mbit/s. Devido

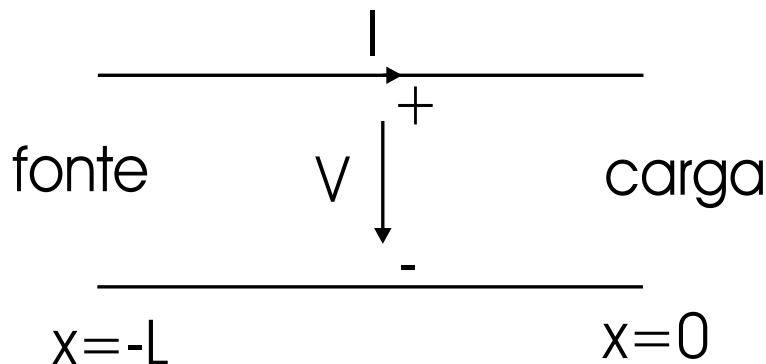


Figura 3.1: linha de transmissão uniforme.

às elevadas frequências e ao comprimento das linhas de transmissão a sua análise reveste-se da particularidade de a tensão e a corrente serem funções não do tempo mas também do espaço. Assim, e considerando a figura 3.1, podemos dizer que a

tensão e a corrente no momento t e no ponto x do eixo colocado ao longo da linha são

$$V(x, \omega) = V(x)e^{j\omega t}, \quad I(x, \omega) = I(x)e^{j\omega t} \quad (3-1.1)$$

onde $V(x)$ e $I(x)$ são dois termos complexos representando a dependência espacial do potencial e corrente eléctrica respectivamente. O potencial e a corrente na linha de transmissão são resultantes da sobreposição de uma onda que se propaga na direcção $x > 0$ e outra na direcção $x < 0$. Podemos então escrever

$$V(x) = V_+e^{-\gamma x} + V_-e^{\gamma x}, \quad I(x) = \frac{1}{Z_0}(V_+e^{-\gamma x} - V_-e^{\gamma x}), \quad (3-1.2)$$

onde os termos V_+ e V_- correspondem às ondas no sentido positivo e negativo respectivamente, γ é chamada a constante de propagação, complexa e dada por $\gamma = \alpha + j\beta$ - α é a atenuação e β é a constante de fase - e finalmente Z_0 é a impedância característica da linha. Substituindo, por exemplo, a parte positiva de (3-1.2) na expressão da tensão de (3-1.1) obtemos que o atraso de fase de um lado ao outro da linha é $\phi = \beta L/2\pi$ e portanto o tempo que demora (à frequência ω) é de

$$t_0 = \frac{\beta}{\omega}L \quad \text{sec} \quad (3-1.3)$$

visto que a distância total é L podemos tirar de (3-1.3) que a velocidade de propagação é $v = \omega/\beta$.

Através das condições de adaptação de impedância podemos dizer que quando uma linha de transmissão se encontra fechada por uma impedância de carga Z_L igual à impedância característica Z_0 então a impedância de entrada é igual à impedância característica e existe uma condição de adaptação em potência.

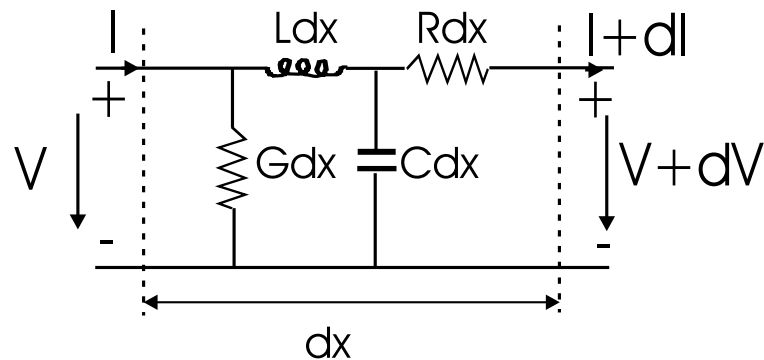


Figura 3.2: modelo paramétrico de uma secção de linha de transmissão.

A forma mais usual de análise de linhas consiste em considerar um modelo físico como o representado na figura 3.2. O interesse deste modelo é de permitir calcular

directamente os valores característicos da linha a partir de constantes físicas, *i.e.*, para o caso da figura 3.2,

$$Z_0 = \sqrt{\frac{R + j\omega L}{G + j\omega C}}, \quad \gamma = \sqrt{(R + j\omega L)(G + j\omega C)}. \quad (3-1.4)$$

No caso de cabos bifilares os parâmetros físicos tomam os seguintes valores típicos:

- capacidade = 0.0515 $\mu\text{F}/\text{km}$, sensivelmente independente da frequência na banda de utilização corrente.
- conductância é extremamente baixa e desprezável.
- inductância = 0.62 mH/km a baixa frequência diminuindo até cerca de 70% deste valor com o aumento da frequência.
- resistência que é aproximadamente proporcional à raiz quadrada da frequência na gama de frequências mais alta, devido ao efeito de pele (tendência para a corrente circular na camada exterior do conductor a alta frequência - *skin effect*).

Como nota adicional podemos referir que existe hoje em dia uma esperança realística de que a introdução de novos materiais super conductores levem ao fabrico de linhas conductoras praticamente sem perdas. Nesse caso ideal teríamos

$$Z_0 = \sqrt{\frac{L}{C}}, \quad \gamma = j\omega\sqrt{LC}, \quad (3-1.5)$$

neste caso a impedância característica é puramente resistiva, por isso fácil de adaptar, e visto que a constante de propagação é imaginária pura o termo de atenuação $\alpha = 0$, como era de esperar. Não existe atenuação ao longo da linha.

3.2 Fibra óptica

A fibra óptica utiliza a luz para transmitir informação. É o meio de transmissão por excelência hoje em dia. O sinal transmitido pode ser do tipo microondas ou no espectro do visível. As frequências são da ordem de, ou superiores a, 10^{14} Hz. Neste caso o canal de transmissão comporta-se como um guia de ondas no qual a construção da fibra é extremamente importante de forma a manter duas camadas (uma interna e outra externa) com índices de refração diferentes de modo a evitar que a onda saia para fora do guia de ondas. É também importante manter o ângulo de reflexão na interface entre as duas zonas sempre inferior a um valor limite θ_m de

modo a que o sinal seja totalmente reflectido e não haja refração evitando assim a perda de energia. Como ilustração do princípio básico vejamos a figura 3.3 onde temos dois meios de propagação 1 e 2, de índices de refração n_1 e n_2 respectivamente. A figura 3.3(a) mostra o caso em que o raio incidente no meio 1 segundo um ângulo

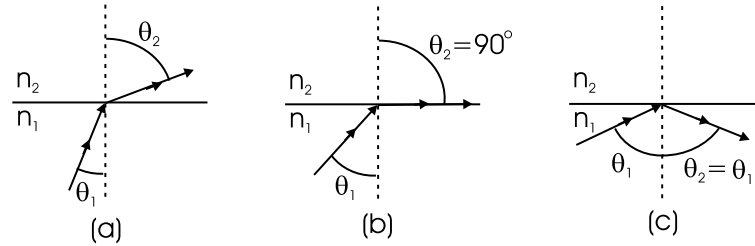


Figura 3.3: ilustração da lei de Snell: a) ângulo de incidência < ao ângulo crítico; b) ângulo de incidência = ao ângulo crítico e c) ângulo de incidência > ao ângulo crítico.

θ_1 é parcialmente transmitido para o meio de 2, segundo um ângulo θ_2 , de acordo com a lei de Snell,

$$\frac{\sin \theta_1}{\sin \theta_2} = \frac{n_2}{n_1} < 1 \quad (3-2.1)$$

Mantendo os mesmos índices de refração mas aumentando o ângulo de incidência chegamos ao valor chamado ângulo de incidência crítica θ_c , para o qual $\theta_2 = \pi/2$ (figura 3.3(b)) e então a partir de (3-2.1) temos

$$\sin \theta_c = \frac{n_2}{n_1}, \quad (3-2.2)$$

e a luz é refractada ao longo da interface entre os dois materiais. Aumentando posteriormente o ângulo de incidência temos reflexão interna total na qual o ângulo de incidência e de reflexão são iguais. É este último modo de propagação que é empregue na fibra óptica.

Existem essencialmente três factores mais importantes que limitam a banda, e por isso a capacidade de transmissão da fibra óptica, que são:

- *atenuação do material*: através de quatro mecanismos: difusão devido a impurezas, absorpção, perdas nos conectores e perdas introduzidas pela curvatura do cabo.
- *dispersão modal*: é devida à diferença de velocidade de propagação dos diferentes modos do guia de ondas. Por outras palavras a energia injectada numa extremidade do guia de ondas não vai chegar à outra extremidade toda ao mesmo tempo. Este facto causa um alargamento dos impulsos transmitidos e por isso uma sobreposição entre símbolos, chamada interferência intersimbólica (ISI - intersymbol interference).

- *dispersão cromática*: é causada pelas diferenças de velocidade de propagação a diferentes comprimentos de onda.

Tem sido feitos enormes progressos ultimamente nos dispositivos de geração e de controlo dos raios luminosos para ataque das fibras ópticas. Atingem actualmente taxas de transmissão típicas de 100 a 1000 GB-km/sec. Este é um campo de investigação intensa hoje em dia.

3.3 Transmissão rádio

Nestes sistemas o sinal é primeiro aplicado a uma antena antes de atravessar o canal de propagação. Inversamente, no receptor, o sinal é primeiro captado por uma antena antes de ser processado para ser retirada a informação útil. O canal de propagação propriamente dito pode ser mesmo assim de vários tipos: a propagação pode ser em linha de vista através da camada baixa da atmosfera; pode ser reflectida nas camadas de ar superiores da atmosfera (ionosfera) e assim o sinal pode ser recebido em pontos não directamente rádio visíveis entre si; ou ainda pode ser através do espaço sem atmosfera como é o caso das comunicações espaciais ou via satélite. As perdas de transmissão das ondas rádio são proporcionais ao logaritmo da distância entre o emissor e o receptor, o que levaria a considerar que as transmissões a longa distância seriam preferivelmente efectuadas através de ondas rádio. Porém esta lei de atenuação só é praticável para emissores em linha de vista o que limita na prática o seu raio de acção.

Vejamos o efeito dos dois factores mais importantes em transmissão via rádio que são a atenuação e o atraso do sinal. Vamos supor que a atenuação é A e a distância é d a uma velocidade de propagação de c provocando um atraso de $\tau = d/c$ segundos. Assim o sinal passabanda recebido escreve-se

$$s(t) = \sqrt{2}A\text{Re}\{u(t - \tau)e^{j\omega_c(t-\tau)}\} \quad (3-3.1)$$

onde $u(t)$ é o sinal transmitido em banda base. Podemos ainda escrever (3-3.1) fazendo sobressair a constante de propagação $\kappa = \omega_c\tau/d$,

$$s(t) = \sqrt{2}\text{Re}\{Au(t - \tau)e^{j\kappa d}e^{j\omega_c t}\} \quad (3-3.2)$$

Podemos então modelizar o canal de transmissão como um filtro de resposta impulsiva

$$h(t) = Ae^{-j\omega\tau}e^{j\kappa d} \quad (3-3.3)$$

mostrando uma dependência da frequência que é linear devido ao atraso τ . Receptores móveis são bastante mais sensíveis a pequenas mudanças em d - que produzem diferenças de fase - do que a variações no atraso.

3.4 Modelos matemáticos de canais de transmissão

De forma a poder estudar e comparar o desempenho de determinados métodos de codificação e modulação, torna-se necessário dispôr de modelos matemáticos representativos dos vários canais de transmissão, que podem ir do caso mais simples do ruído aditivo até ao mais complexo do filtro linear variante no tempo.

3.4.1 Canal de ruído aditivo

Á parte o caso em que o canal não introduz nenhuma alteração no sinal, este é sem dúvida o caso mais simples de um canal realístico no qual o sinal emitido $s(t)$, chega ao receptor simplesmente adicionado com um termo de ruído $w(t)$, tal que o sinal recebido $r(t)$, se pode escrever

$$r(t) = s(t) + w(t), \quad (3-4.1)$$

onde o termo de ruído é uma realização de um processo estocástico devido a ruído de origem electrónica no sistema de emissão/recepção ou devido a interferências no meio físico de propagação do sinal. É frequente, e não desprovido de sentido prático, considerar que a sequência $w(t)$ é branca, que segue uma lei conjuntamente Gaussiana e é independente do sinal emitido $s(t)$. Uma atenuação e um atraso de propagação podem ser facilmente considerados, sem que seja introduzida distorção, através de

$$r(t) = \alpha s(t - t_0) + w(t), \quad (3-4.2)$$

onde α e t_0 são constantes.

3.4.2 Canal de filtro linear invariante

Este caso é ligeiramente mais complexo do que o precedente, englobando-o. Aqui o sinal recebido é suposto ser a soma de um termo de ruído aditivo (como no caso anterior) e um termo sem ruído igual á resposta de um filtro linear invariante ao sinal emitido $s(t)$. Assim, temos que o sinal recebido $r(t)$, se escreve

$$\begin{aligned} r(t) &= g(t) * s(t) + w(t), \\ &= x(t) + w(t), \end{aligned} \quad (3-4.3)$$

onde $g(t)$ é a resposta impulsiva do canal considerado. Podemos então escrever

$$x(t) = \int g(\tau) s(t - \tau) d\tau. \quad (3-4.4)$$

3.4.3 Canal de filtro linear variante no tempo

A representação do canal de transmissão através de um filtro linear variante no tempo, representa um grau de complexidade acrescida em relação ao caso anterior, mas que é muitas vezes justificado na prática. A única diferença é que a componente sem ruído do sinal recebido se escreve

$$x(t) = \int g(\tau; t)s(t - \tau)d\tau, \quad (3-4.5)$$

onde neste caso a resposta impulsiva do canal $g(t)$ é variante no tempo. O exemplo mais conhecido para um canal linear variante no tempo é o dos telefones celulares e da propagação do sinal acústico submarino em águas pouco profundas. Nestes casos o sinal chega ao receptor móvel através de um número elevado de canais de propagação (propagação multicanal) cada um com um atraso e uma atenuação específicos. Normalmente a atenuação de cada canal é também variante no tempo.

4 Quantidade de informação, quantificação e codificação

Tivemos oportunidade de abordar no capítulo 2.1.3 a questão da passagem do sinal contínuo $s(t)$ para a sua versão discreta no tempo $s(k)$. Em particular, preocupámo-nos em deduzir um processo que garantisse duas coisas: 1) que a representação discreta ocupasse o mínimo de banda possível e 2) que permitisse uma reconstrução fiel e sem ambiguidade do sinal original. Por outras palavras, uma reconstrução fiel significa uma reconstrução com o mínimo de distorção. Normalmente diz-se que a distorção é mínima, neste caso zero, quando a quantidade de informação é preservada. Isto é quando a informação contida no sinal $s(t)$ se encontra integralmente no sinal $s(k)$.

É facilmente compreensível que a noção de quantidade de informação seja uma questão central em problemas de telecomunicações. O objectivo primário de um sistema de comunicação não é mais do que transmitir informação de um ponto para o outro. Foi Claude Shannon em 1940 que, pela primeira vez, introduziu a possibilidade de quantificar a quantidade de informação, e assim revolucionou a nossa forma de ver a problemática das comunicações.

Porém não basta que uma determinada mensagem seja fielmente discretizada no tempo, é também necessário que ela seja discretizada na sua amplitude. Isto é, $s(t)$ é à partida uma função real de variável real t . Uma vez que se procedeu à discretização da variável tempo t , ela passou a ser uma função real de variável discreta k , que toma valores inteiros. Teoricamente uma função real pode tomar um número infinito de valores o que, utilizando um sistema binário, necessitaria na prática um número infinito de bits para ser representado. O processo que permite reduzir o número de níveis que pode tomar a amplitude de uma determinada mensagem chama-se quantificação. A quantificação é um processo que introduz inevitavelmente alguma distorção.

Finalmente, quando a mensagem que se pretende transmitir se encontra discretizada, seja no tempo seja na amplitude, trata-se de determinar qual o melhor modo de representar os níveis obtidos após discretização de forma a minimizar, tanto quanto possível, a quantidade de dados a transmitir, sem que no entanto haja perda de informação. A este último processo chama-se codificação sem ruído ou compressão de dados. Talvez esta última etapa não seja muito clara e por isso vamos dar um exemplo simples. Vamos imaginar que uma determinada mensagem é quantificada entre -10 e +10 volts utilizando 1024 níveis discretos. Admitindo que o sinal é em geral de média nula, é provável que uma estatística demonstre que os valores dos níveis centrais (em torno a 0) são mais prováveis do que os níveis perto dos -10 ou dos +10 volts. Compreende-se então facilmente que seria vantajoso para o sistema de comunicações que os níveis centrais fossem codificados com menos bits, enquanto

os níveis menos prováveis (que ocorrem menos frequentemente) fossem codificados com mais bits. Este processo de atribuição de códigos mais curtos a símbolos mais prováveis e de códigos mais longos a símbolos menos prováveis, consiste um exemplo simples da utilização eficiente da codificação para a diminuição do volume de dados a transmitir, aumentando assim a eficiência do sistema de transmissão. Trata-se de alguma forma de uma compressão de dados, cujo exemplo mais familiar de todos é o algoritmo de Lempel-Ziv universalmente utilizado nos programas de “zipagem” em informática.

4.1 Quantidade de informação e entropia

A entropia é uma forma de medir a quantidade de informação. Intuitivamente uma mensagem tem tanta mais informação quanto maior for o seu grau de aleatoriedade ou imprevisibilidade. Por exemplo, se dissermos: “o sol nasceu esta manhã”. Temos aqui uma mensagem com uma certa quantidade de informação. Agora se dissermos: “houve um terramoto em Lisboa.” Qual das duas frases tem mais informação? Intuitivamente seremos levados a dizer que a segunda tem mais informação que a primeira. Isto, porque a segunda mensagem era mais imprevisível, ou seja representa um acontecimento que ocorre menos vezes que o da primeira mensagem. A entropia tenta traduzir uma forma de quantificar este conceito. Assim, por exemplo, a entropia ou quantidade de informação da variável aleatória x de probabilidade $p_X(x)$, é definida por

$$H(x) = E[-\log_2 p_X(x)] = - \sum_{x \in \Omega_x} p_X(x) \log_2 p_X(x). \quad (4-1.1)$$

Nesta equação percebemos o porquê da densidade de probabilidade e do sinal de menos, pois a entropia deverá ser tanto maior quanto menos provável for o sinal, e também percebemos o porquê da esperança matemática $E[.]$, que realiza uma “média” sobre todas as tiragens possíveis. Mas de onde vem o \log_2 ? A introdução do logaritmo está ligada com o facto de que se considerarmos dois processos estocásticos independentes X e Y , a informação do acontecimento conjunto de X e de Y deverá ser simplesmente a soma das quantidades de informação de um e do outro (o logaritmo do produto é a soma dos logaritmos!). Assim deverá haver um logaritmo tal que

$$H(x, y) = E[-\log_2 p_{XY}(x, y)] \quad (4-1.2)$$

$$= E[-\log_2 p_X(x) - \log_2 p_Y(y)] \quad (4-1.3)$$

$$= H(x) + H(Y), \quad (4-1.4)$$

onde $p_{XY}(x, y)$ não é mais do que a densidade de probabilidade conjunta das va's X e Y , tal que, visto que as va's são independentes, $p_{XY}(x, y) = p_X(x)p_Y(y)$. Para o caso do alfabeto binário, formado de 0 e 1's, com probabilidades p e $1 - p$ a entropia total é dada por

$$H_b(p) = -p \log_2 p - (1 - p) \log_2(1 - p) \quad (4-1.5)$$

que se encontra representada na figura 4.1.

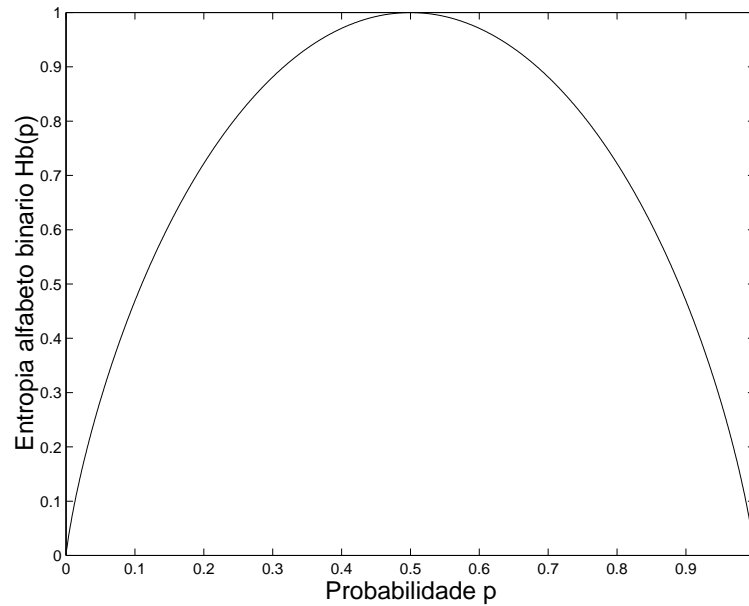


Figura 4.1: entropia do alfabeto binário.

4.2 Quantificação

Quantificação é o processo de representação de uma fonte de dados analógicos num conjunto finito de níveis. Como já dissemos, este processo introduz necessariamente alguma distorção e perda de informação, o que é inevitável. De uma forma geral a quantificação pode ser *uniforme* ou *não uniforme*. Na quantificação uniforme o passo de quantificação, i.e., a diferença entre dois níveis contíguos é constante entre o valor mínimo e o valor máximo. No caso de quantificadores não uniformes, pelo seu lado, o nível de quantificação não é constante. Na análise da quantificação pode-se representar o sinal a quantificar por uma variável aleatória (v.a.) X , de lei $p_X(x)$. O processo de quantificação consiste na divisão do domínio de X em N intervalos contíguos onde cada intervalo de quantificação R_n é representado por um valor \hat{x}_n . Assim cada vez que a v.a. X , de valor x , à entrada do quantificador pertence ao intervalo R_n , a saída do quantificador toma o valor \hat{x}_n . O erro quadrático médio

$$\begin{aligned}\epsilon &= E[(x - \hat{x}_n)^2] \\ &= \sum_{n=1}^N \int (x - \hat{x}_n)^2 p_X(x) dx,\end{aligned}\tag{4-2.1}$$

caracteriza, em média, o erro quadrático de quantificação cometido ao substituir o valor real x pelo seu valor quantificado \hat{x}_n . O valor (4-2.1) toma o nome de ruído de

quantificação. Assim, a relação sinal-ruído de quantificação escreve-se (SQNR=signal to quantization noise ratio).

$$\text{SQNR}_{\text{dB}} = 10 \log_{10} \frac{E[X^2]}{\epsilon}. \quad (4-2.2)$$

onde $E[X^2]$ representa a esperança matemática do quadrado da v.a. X , i.e., a sua variância se for de média nula.

4.2.1 Quantificação uniforme

No processo de quantificação uniforme o domínio de discretização é dividido em intervalos iguais, i.e., $\hat{x}_{n+1} - \hat{x}_n = \Delta$ é uma constante $\forall n$. Se o domínio a discretizar não for limitado, então o primeiro e o último intervalos de quantificação são infinitos, i.e., as regiões de discretização serão dadas por

$$\begin{aligned} R_1 &=] - \infty, a] \\ R_2 &=]a, a + \Delta] \\ R_3 &=]a + \Delta, a + 2\Delta] \\ &\vdots \end{aligned} \quad (4-2.3)$$

$$R_N =]a + (N - 2)\Delta, \infty[\quad (4-2.4)$$

demonstra-se facilmente que o nível a associar com cada intervalo de quantificação é dado pelo valor médio de cada intervalo, assim

$$\hat{x}_n = a + (n - 2)\Delta + \frac{\Delta}{2}, \quad (4-2.5)$$

este ponto é chamado a centróide do intervalo e não é mais do que $\hat{x}_n = E[X|X \in R_n]$. Será fácil determinar que a função de quantificação $Q(x) = \hat{x}_n, \quad n \in [1, N]$ é uma função em escada, dependente dos dois parâmetros a e Δ .

4.2.2 Quantificação não uniforme

Neste caso o intervalo de quantificação não é constante. É óbvio que uma zona de valores onde o sinal é mais provável deverá ter mais níveis do que uma região onde o sinal é pouco provável. Assim, este método de quantificação permite normalmente obter um maior desempenho que o caso uniforme. Prova-se com efeito que os intervalos e níveis de quantificação são dados de forma óptima através das condições de Lloyd-Max que se escrevem

$$\begin{aligned} \hat{x}_n &= \frac{\int_{a_{n-1}}^{a_n} x p_X(x) dx}{\int_{a_{n-1}}^{a_n} p_X(x) dx} \\ a_n &= \frac{\hat{x}_{n-1} + \hat{x}_n}{2} \end{aligned} \quad (4-2.6)$$

de onde podemos concluir que os níveis de quantificação óptimos são dados pelas centroides e os limites de quantificação de cada região são dados pelos pontos médios entre cada par de centroides. A resolução do sistema de equações (4-2.6) faz-se de modo iterativo partindo de um conjunto inicial de centroides e calculando iterativamente o ruído de distorção, até que não haja uma diminuição significativa dessa distorção entre um ponto e o seguinte.

4.3 Codificação

Já vimos como colocar uma mensagem sob a forma de uma sequência de símbolos que forma uma representação discreta da informação emitida pela fonte. Vamos agora tratar do problema de como codificar de forma óptima esta sequência de símbolos numa sequência de dígitos binários. Tendo em conta que o alfabeto utilizado pela fonte é geralmente finito, o problema da codificação torna-se um problema relativamente fácil. No entanto isto só é verdade quando os símbolos são estatisticamente independentes entre si, i.e., quando o sistema não tem memória - diz-se nesse caso que temos uma *fonte discreta sem memória*. Desfortunadamente, na prática, poucas fontes de mensagens podem ser consideradas como sendo *fontes discretas sem memória*.

O exemplo clássico consiste em considerar uma fonte discreta sem memória que produz um símbolo cada τ segundos. Cada um dos símbolos é selecionado de um alfabeto finito com L símbolos $x_n; n = 1, \dots, L$, com a probabilidade $p(x_n)$. Assim a entropia da fonte, de acordo com (4-1.1), é dada por

$$H(x) = - \sum_{n=1}^L p(x_n) \log_2 p(x_n). \quad (4-3.1)$$

Se os símbolos forem igualmente prováveis temos que $p(x_n) = 1/L$ e então

$$\begin{aligned} H(x) &= -\frac{1}{L} \sum_{n=1}^L \log_2 \frac{1}{L} \\ &= \log_2 L, \end{aligned} \quad (4-3.2)$$

o que representa o valor máximo que pode tomar $H(x)$ em (4-3.1), e que é o número de bits médio necessário para representar cada símbolo do alfabeto. A taxa de transmissão é então de $H(x)/\tau$ bits/s. Existem essencialmente dois grandes grupos de métodos de codificação para fontes discretas sem memória, que são: codificação com palavras de comprimento fixo e codificação com palavras de comprimento variável.

4.3.1 Com palavras de comprimento fixo

Neste caso utiliza-se um número de bits R fixo para cada símbolo. De acordo com este esquema de codificação, a cada símbolo corresponderá um único conjunto de R bits. Dado que temos L símbolos e, admitindo que L é uma potência de 2, o número de bits necessário é dado por $R = \log_2 L$. Se L não for uma potência de 2 então o número de bits necessário será $R = \text{int}[\log_2 L] + 1$, onde int significa “o maior inteiro inferior a”. É óbvio que se $H(x) \leq \log_2 L$ teremos que $R \geq H(x)$. Na prática um bom esquema de codificação é aquele que se aproxima o mais possível da entropia, e portanto $H(x)/R$ é uma medida da eficiência do esquema de codificação. Existem métodos alternativos que consistem em codificar não cada símbolo individualmente mas sim sequências de símbolos (um exemplo típico é um código fonte no qual as palavras *while do*, etc se encontram frequentemente repetidas). Assim pode-se melhorar a eficiência do esquema de codificação aumentando o número de símbolos por grupo tanto quanto se pretender.

Alternativamente, existem esquemas nos quais a relação entre a representação em bits de cada símbolo não é única. Isto significa que no alfabeto podem coexistir um conjunto de símbolos cuja representação é única (os mais frequentes) e um outro conjunto de símbolos representados pelo mesmo conjunto de bits (os muito pouco frequentes). É claro que este esquema de codificação não permite a recuperação de toda a informação emitida pela fonte no momento de descodificação. Trata-se aqui de uma codificação que introduz ruído de codificação e portanto perda de informação. O problema na prática é de controlar o ruído de quantificação dentro de níveis aceitáveis obtendo em troca um forte aumento da compactação da informação que se traduz numa eficiência alta e portanto uma alta taxa de transmissão efectiva.

4.3.2 Com palavras de comprimento variável

Quando os símbolos emitidos pela fonte não são equiprováveis torna-se óbvio que o esquema de codificação mais apropriado deverá ter códigos de comprimento variável: os símbolos mais frequentes deverão ter códigos com menos bits e vice-versa. O exemplo mais conhecido é o código Morse que já na época utilizava códigos mais curtos para as letras mais frequentes do alfabeto. Um problema óbvio desde logo é que as letras mais frequentes por exemplo em inglês não serão sempre as mais frequentes em português o que leva inevitavelmente a perdas de eficiência quando de mudanças de língua. Um problema que se coloca na prática quando da codificação com palavras de comprimento variável é que por vezes não podemos saber imediatamente no receptor se uma dada sequência de bits forma vários símbolos codificados com códigos curtos ou símbolos codificados com códigos longos, dado que utilizando dígitos binários, utilizar o ‘1’ para o A, ‘00’ para o B, ‘01’ para o C e ‘10’ para o D, a sequência ‘001001’ não nos permite determinar unicamente quais os símbolos emitidos. Clara-

mente temos B como primeiro símbolo mas em seguida não sabemos se é ABA ou CB. Na prática temos que ter a certeza que o código escolhido não contém ambiguidades e pode ser unicamente e instantaneamente decodificado. Desde já podemos notar que existe uma condição de base que é chamada a *Condição do prefixo*: e que consiste em afirmar que nenhum código deverá ser o prefixo de um outro código. Pode-se efectivamente demonstrar que códigos com palavras de comprimento variável que satisfazem a condição do prefixo são eficientes para qualquer fonte discreta sem memória com símbolos não equiprováveis. O algoritmo de Huffman permite construir tais códigos e é muito importante na prática.

Algoritmo de codificação de Huffman

O método proposto por Huffman em 1952 baseia-se no princípio de atribuir de forma progressiva os códigos mais longos aos símbolos menos prováveis até aos símbolos mais prováveis, evitando os prefixos. Desde logo se pode constatar que para construir um esquema de codificação com o algoritmos de Huffman é necessário dispôr das probabilidades de cada símbolo. O procedimento inicia-se concatenando os dois símbolos menos prováveis construindo um novo símbolo cuja probabilidade é igual à soma das probabilidades dos dois símbolos iniciais. Este processo prossegue até ficarmos com um único símbolo de probabilidade 1. Obtem-se assim uma árvore contendo todas as combinações dos símbolos existentes. Começa-se depois pela raiz da árvore alocando os dígitos ‘0’ e ‘1’ aos dois símbolos mais prováveis (e que assim terão o código mais curto); em seguida os dois símbolos que formam, por exemplo o símbolo codificado com o ‘0’ serão codificados com ‘00’ e ‘01’, da mesma forma para o outro ramo com ‘10’ e ‘11’ e assim sucessivamente aumentando o número de dígitos do código à medida que se diminui a probabilidade do símbolo e garantido que a condição do prefixo é satisfeita.

Voltamos agora ao caso geral no qual a fonte não é sem memória, i.e., o caso em que os símbolos emitidos não são estatisticamente independentes uns dos outros, porém consideraremos sempre o caso em que a sequência de símbolos será considerada estacionária. Ou seja, o processo aleatório discreto associado à fonte não é branco mas é estacionário. Aqui o problema é singularmente mais complicado dada a dependência estatística entre símbolos, que não nos permite determinar a entropia (e portanto o número mínimo de bits necessário), de forma directa. Felizmente pode-se definir a entropia de um alfabeto de uma fonte estacionária através da definição da entropia de um bloco de símbolos. Assim, seja a entropia do bloco de símbolos $\{X_1, X_2, \dots, X_k\}$,

$$H(X_1, X_2, \dots, X_k) = \sum_{i=1}^k H(X_i | X_1, X_2, \dots, X_{i-1}), \quad (4-3.3)$$

onde $H(X_i | X_1, X_2, \dots, X_{i-1})$ é a entropia condicional do símbolo X_i dados os símbolos anteriores $X_n; n = 1, \dots, i - 1$. A entropia por símbolo dentro do conjunto de k

símbolos, define-se como sendo

$$H_k(X) = \frac{1}{k}H(X_1, X_2, \dots, X_k). \quad (4-3.4)$$

Finalmente a entropia do alfabeto escreve-se como sendo a entropia do bloco k quando $k \rightarrow \infty$,

$$H(X) = \lim_{k \rightarrow \infty} H_k(X) = \lim_{k \rightarrow \infty} \frac{1}{k}H(X_1, X_2, \dots, X_k). \quad (4-3.5)$$

O papel da definição da noção de entropia no caso estacionário é essencial pois permite estabelecer uma métrica da eficiência da codificação. A partir daqui prova-se que uma codificação eficiente de uma fonte discreta estacionária com memória é possível através da representação de largos blocos de símbolos por códigos. Para cada um dos blocos pode (e deve) ser utilizado o algoritmo de Huffman com código de palavras de comprimento variável. De notar no entanto que para que isso seja possível será necessário conhecer as densidades de probabilidade conjuntas dos símbolos contidos no mesmo bloco, o que nem sempre acontece de forma directa e requiere frequentemente na prática uma fase de amostragem para obter uma estimação empírica dessas probabilidades. Para obviar este problema da necessidade da estimação das probabilidades a priori foi proposto o algoritmo de Lempel-Ziv.

Algoritmo de Lempel-Ziv

Este algoritmo é independente das propriedades estatísticas dos dados e por isso tem um carácter universal. Este algoritmo procede separando a sequência emitida pela fonte em blocos de comprimento variável, chamadas frases. Uma nova frase é criada cada vez que ela difere no seu último símbolo de uma outra frase formada anteriormente. As frases são colocadas num dicionário juntamente com o respectivo índice de início de frase na mensagem. Quando toda a mensagem foi lida, atribui-se sequencialmente um código a cada uma das frases formado pela localização da frase na mensagem acrescentada da última letra em que cada mensagem difere das outras já classificadas no dicionário. Esta etapa de codificação é inicializada começando com o código '0000' (no número de bits necessário para o comprimento do dicionário). No decodificador constrói-se a tabela a partir dos códigos recebidos procedendo ao contrário, separando o código em localização e última letra de forma ordenada. A partir da tabela reconstrói-se a mensagem. Trata-se de um processo não linear cuja análise ultrapassa largamente o objectivo deste capítulo introdutivo à codificação. Mencione-se no entanto que este algoritmo é hoje em dia o mais utilizado em aplicativos de compressão e descompressão de dados tais como os programas `zip` e `unzip` tanto em Unix como em DOS.

4.3.3 Codificação de sinais analógicos

Até agora temos descrito técnicas de codificação para fontes de sinal discretas, ou para fontes de sinal analógicas discretizadas e quantificadas. Existem no entanto

técnicas de codificação utilizadas para fontes de sinal analógicas como, por exemplo, a Modulação por códigos de pulsos (*Pulse Code Modulation -PCM*) e as suas variantes tal como a PCM diferencial (DPCM) ou PCM diferencial adaptativa (ADPCM). Vamos agora definir estes tipos de codificação.

Modulação por Códigos de Pulso (Pulse Code Modulation - PCM)

Existem dois tipos de PCM: uniforme e não uniforme. Na PCM uniforme o sinal analógico é primeiramente amostrado a uma taxa superior a Nyquist, i.e., se W for a banda (em Hz) do sinal analógico, a taxa de amostragem deverá ser $f_s \geq 2W$. Em seguida o sinal já discreto no tempo é quantificado utilizando 2^R níveis discretos entre o valor mínimo e máximo do sinal de entrada $x(t) \in [-x_{\max}, x_{\max}]$. Assim o intervalo de quantificação será $\Delta = 2x_{\max}/2^R$. Como em qualquer processo de quantificação vai existir um erro (ou ruído) de quantificação devido à diferença entre o valor real do sinal no instante k , $x(k)$ e o seu valor quantificado $\hat{x}(k)$, que pode ser estimado considerando que, se o número de níveis $N = 2^R$ é suficientemente elevado, o sinal amostrado pode-se representar como

$$\hat{x}(k) = x(k) + q(k), \quad (4-3.6)$$

onde $q(k)$ é o ruído de quantificação, suposto uniforme no intervalo $[-\Delta/2, \Delta/2]$. Ora é sabido que a potência do ruído pode ser expressa pela sua variância, o que para uma v.a. uniforme é

$$\begin{aligned} \sigma^2 &= \frac{\Delta^2}{12}, \\ &= \frac{x_{\max}^2}{3 \times 4^R}, \end{aligned} \quad (4-3.7)$$

$$(4-3.8)$$

de onde podemos calcular a relação sinal-ruído de quantificação (SQNR) em dB, como sendo

$$\begin{aligned} \text{SQNR}_{\text{dB}} &= 10 \log_{10} \left[3 \times 4^R \frac{E[X^2]}{x_{\max}^2} \right] \\ &= 10 \log_{10} [3 \times 4^R E[\bar{X}^2]] \\ &\approx 4.8 + 6R + 10 \log_{10} E[\bar{X}^2], \end{aligned} \quad (4-3.9)$$

onde $\bar{X} = X/x_{\max}$ é o valor normalizado do sinal $x(k)$. A taxa de transmissão do sinal codificado é (no mínimo) $f_s R$, o que significa que na prática é quase sempre superior a $1.5 f_s R$.

A PCM não uniforme difere apenas pelo facto de que antes de ser quantificado, o sinal é filtrado por um elemento não linear destinado a reduzir a sua gama de valores. Na prática isto corresponde a ter mais níveis nos valores em torno a 0 e menos níveis

nos valores extremos. Depois desta filtragem em amplitude não linear é aplicada um PCM uniforme. Na recepção, depois de uma decodificação PCM uniforme o sinal deve ser filtrado com um filtro inverso daquele utilizado na transmissão de forma a colocar o sinal na sua gama de valores original. Nas aulas práticas veremos alguns exemplos de filtragem não linear para PCM não uniforme.

PCM diferencial (*Differential PCM - DPCM*)

Em PCM cada uma das amostras é quantificada e codificada nos 2^R bits independentemente das outras amostras anteriores do sinal. Ora na prática, se a taxa de amostragem f_s é suficientemente alta, a amplitude do sinal de entrada é suposta não variar muita numa amostra para a seguinte, o que pressupõe a ideia de que em vez de codificar o valor absoluto do sinal a codificação do seu valor em relação à amostra anterior permitiria utilizar muito menos bits e portanto uma menor taxa de transmissão para enviar a mesma quantidade de informação. Esta é a ideia de base do PCM diferencial (DPCM): trata-se de codificar não o valor absoluto do sinal mas apenas a sua variação em relação à amostra anterior. Um dos problemas mais óbvios com a aplicação directa desta simples ideia é de que os erros devidos ao ruído tem tendência para se adicionar ao longo do tempo. Por essa razão surge a ideia de não limitar a dependência da evolução do sinal apenas à última amostra, mas sim às, digamos, p últimas amostras. Assim, constrói-se um modelo de dados recursivo de forma a escrever o sinal amostrado predicto no instante k como

$$\tilde{x}(k) = \sum_{i=1}^p a_i x(k-i), \quad (4-3.10)$$

o que é uma combinação linear ponderada das p amostras anteriores, e onde os coeficientes de ponderação a_i são determinados de forma a minimizar o erro quadrático médio $\epsilon^2(k)$ entre o sinal predicto e o sinal real

$$E[\epsilon^2(k)] = E[(x_k - \sum_{i=1}^p x_{k-i})^2]. \quad (4-3.11)$$

Existem vários métodos bem conhecidos para estimar os coeficientes a_i , entre os quais os famosos algoritmos de Levison e Durbin (1959). Na prática o que acaba por ter que ser transmitido é o erro entre o sinal predicto e o sinal real o que requiere, no lado do receptor, que o mesmo conjunto de coeficientes a_i seja utilizado para voltar a reconstruir o sinal $x(k)$ a partir de \tilde{x}_k e das amostras anterior estimadas do sinal. Prova-se que neste algoritmo os erros não se acumulam ao longo do tempo.

PCM diferencial adaptativo (*Adaptive DPCM - ADPCM*)

Uma das suposições que estão na base do DPCM é que o sinal não muda muito de uma amostra para a seguinte, o que permite obter um bom estimador do seu valor no instante k conhecendo as amostras nos p instantes anteriores. Quanto mais elevado for

o valor de p maior é a quantidade de informação reunida no algoritmo para determinar $x(k)$ e normalmente mais estável e preciso será o algoritmo, o que na prática se traduz por uma baixa taxa de erro de transmissão. Porém, não devemos esquecer que na base do método se encontra a suposição de invariabilidade (ou fraca variabilidade temporal) do sinal o que, na presença de ruído, se traduz por estacionaridade. A pergunta que se pode colocar é: então e quando o sinal não é estacionário? A resposta a esta pergunta encontra-se no DPCM adaptativo (ADPCM). Na prática a não estacionaridade é difícil de ter em conta, mas também existem poucos processos de interesse que sejam absolutamente não estacionários e portanto apenas nos interessa o caso da quasi-não estacionaridade, onde a média e a variância variam lentamente com o tempo. Uma das formas para ter em conta esta quasi-não estacionaridade é utilizar um quantificador DPCM uniforme no qual o intervalo de quantificação varia de acordo com a variância local das amostras do sinal de entrada. Por variância local entende-se a variância estimada numa janela temporal em torno à amostra actual. Assim pode-se fazer variar o intervalo de amostragem e portanto a dinâmica do sinal à entrada do quantificador de acordo com as variações lentas do sinal de entrada.

Modulação Delta (*Delta Modulation - DM*)

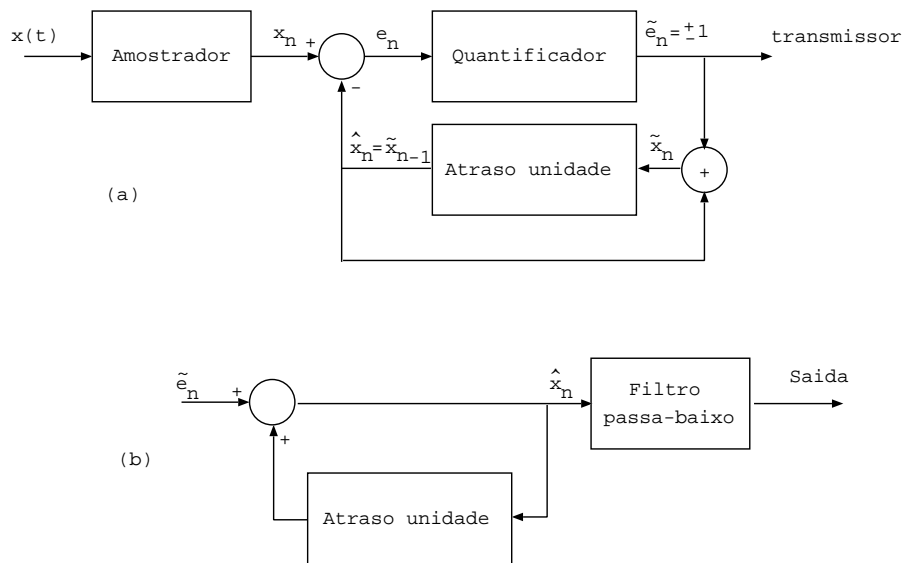


Figura 4.2: codificação por modulação delta (DM): sistema codificador (a) e sistema decodificador (b).

A modulação delta (DM) é uma variante simplificada do DPCM e que se encontra representada na figura 4.2. Neste codificador simplificado, o sinal é modelado utilizando um predictor fixo de ordem um ($p = 1$) do tipo (4-3.10) com $a_1 = 1$ e quantificado apenas com um bit (2 níveis). Na prática isto é equivalente a dizer que o valor predicto do sinal é apenas o valor no instante anterior mais o ruído de quantificação. O resultado é que o sinal se encontra aproximado por uma função em

escada com um determinado intervalo de quantificação Δ . Para ser efectiva, a DM geralmente requiere uma taxa de amostragem bem superior a Nyquist, da ordem de 5 vezes a taxa mínima, e que o sinal a emitir evolua lentamente. Mesmo assim existem dois tipos de distorção frequentes no receptor, que são: o erro de “slope-overload” e o erro de ruído granular. Estes dois tipos de distorção encontram-se exemplificados na figura 4.3 e resultam, como se pode verificar, de uma demasiado rápida variação da entrada para o intervalo de quantificação, que não consegue acompanhar o sinal, e de uma variação demasiado lenta da entrada que resulta numa constante oscilação em torno a um valor que varia demasiado lentamente. Estes dois erros impõem re-

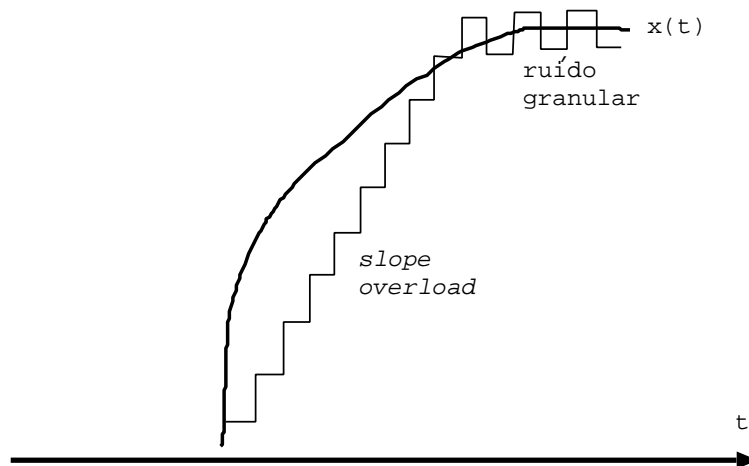


Figura 4.3: distorção de *slope overload* e ruído granular, devida ao desajuste do intervalo de quantificação.

strições contraditórias no intervalo de quantificação Δ que num caso deverá ser maior e no outro caso menor. Existe uma DM com intervalo de quantificação a intervalo variável, na qual Δ é fixado de forma adaptativa de acordo com a maior ou menor taxa de variação do sinal de entrada de forma minimizar simultâneamente os dois tipos de distorção.

4.3.4 Codificação em sub-bandas

Devido a estes problemas de sinais com requisitos contraditórios, uns que variam rápido demais e outros lentos demais, colocou-se a ideia de filtrar o sinal a emitir em várias sub-bandas de frequência, podendo assim, eventualmente separar componentes do sinal com maior dinâmica de outros com menor dinâmica, e utilizando mais bits para um e menos para outro. Isto implica uma recombinação do sinal à chegada ao receptor. Demonstrou-se na prática que a codificação em sub-bandas é muito efectiva para a codificação da voz. Assumindo que a voz se encontra entre 0 e 3200 Hz, é frequente definir 4 sub-bandas 0-400, 400-800, 800-1600 e 1600-3200 Hz. Note-se

que cada filtro tem uma banda que é o dobro do anterior e são por isso chamados filtros em oitavos de banda. A voz tem geralmente uma dinâmica mais elevada e os erros de quantificação são mais perceptíveis ao ouvido também nessa banda, o que implica que mais bits deverão ser usados para codificar o sinal nas primeira banda e progressivamente menos nas seguintes. Geralmente utiliza-se ADPCM em cada sub-banda.

4.3.5 Codificação baseada em modelos do sinal

Convém ainda referir, para completar este capítulo, que existe uma outra classe de métodos que tenta resolver o problema da codificação através de um método completamente diferente que é utilizando modelos do sinal. Neste tipo de métodos o que é transmitido deixa de ser o sinal propriamente dito (ou suas componentes) mas sim os coeficientes de um modelo ajustado no sinal. Para certos tipos de sinais, que podem ser correctamente representados pelo modelo utilizado, o ganho em termos de taxa de transmissão e número de bits utilizado pode ser muito superior a qualquer um dos outros métodos de codificação utilizados anteriormente. Um erro de quantificação mínimo é obtido através de um maior ou menor ajuste do modelo e que se traduz por um maior ou menor número de coeficientes. Obviamente que, se o sinal for dificilmente representado pelo modelo, o número de coeficientes necessários para uma representação correcta torna-se muito elevado e devido a limitações práticas no número máximo de coeficientes, a distorção tende a aumentar. Não nos vamos alongar aqui neste tipo de métodos, basta dizer que o mais utilizado se chama codificação de predição linear (*Linear Prediction Coding - LPC*), e que passa pela utilização de um modelo auto-regressivo (AR) semelhante ao representado na equação (4-3.10).

5 Modulação analógica de onda sinusoidal

A modulação de portadora sinusoidal compreende a modulação de amplitude, de fase e de frequência, que constituem uma família de métodos de modulação nos quais a amplitude, a fase ou a frequência de uma senoide de frequência central pré-determinada é alterada em função do sinal modulador. A modulação de amplitude (AM) é caracterizada por uma relativa simplicidade de representação e uma fraca necessidade de banda passante. Por outro lado a sua eficiência em termos de potência é bastante baixa quando comparada com os métodos de modulação angular (fase e frequência). Os métodos de modulação AM são ainda hoje frequentemente empregues em difusão rádio e TV, comunicações ponto a ponto (SSB) e multiplexagem em telefonia. Os métodos de modulação de fase e frequência são mais difíceis de implementar mas bastante mais eficientes e imunes ao ruído e daí uma qualidade de recepção bastante superior.

5.1 Modulação de amplitude com duas bandas laterais (AM-DSB)

Uma portadora de frequência central f_c e amplitude A_c é modulada em amplitude por um sinal $m(t)$ quando se escreve

$$s(t) = [1 + a\bar{m}(t)]c(t) \quad (5-1.1)$$

onde $m(t)$ é a mensagem a transmitir, passa-baixo, dentro da banda $-W, W$ e $c(t) = A_c \cos(2\pi f_c t)$ é a portadora onde se supõe que $f_c > W/2$ e $0 \leq a \leq 1$ é um coeficiente real denominado índice de modulação. O sinal $\bar{m}(t)$ é a mensagem $m(t)$ normalizada, *i.e.*, tal que

$$\bar{m}(t) = \frac{m(t)}{\max |m(t)|} \quad (5-1.2)$$

Óbviamente (como aliás já vimos no caso da amostragem - capítulo 2.1) este tipo de modulação resulta no domínio da frequência num espectro que se escreve

$$S(f) = \frac{A_c}{2} [a\bar{M}(f - f_c) + \delta(f - f_c) + a\bar{M}(f + f_c) + \delta(f + f_c)], \quad (5-1.3)$$

onde o espectro do sinal modulador (normalizado) foi deslocado de $-f_c$ e $+f_c$, frequências às quais encontramos igualmente um impulso correspondente à portadora. Todo o espectro foi multiplicado por $1/2$ devido ao princípio de conservação de energia. A banda ocupada pelo sinal modulador é de $2W$.

Exemplo: considere o sinal

$$m(t) = \begin{cases} 1 & 0 \leq t < t_0/3 \\ -2 & t_0/3 \leq t < 2t_0/3 \\ 0 & \text{para todo outro } t \end{cases} \quad (5-1.4)$$

com $t_0 = 0.15$ e que modula uma portadora de frequência $f_c = 250$ Hz, com $A_c = 2$ e um índice de modulação $a = 0.8$.

- determine a expressão do sinal modulador normalizado $\bar{m}(t)$
- calcule a expressão do sinal modulado $s(t)$
- determine o espectro de $\bar{m}(t)$ e de $s(t)$

O sinal modulador normalizado escreve-se facilmente, tendo em conta que

$$\max |m(t)| = 2$$

o que implica que

$$\bar{m}(t) = \begin{cases} 1/2 & 0 \leq t < t_0/3 \\ -1 & t_0/3 \leq t < 2t_0/3 \\ 0 & \text{para todo outro } t \end{cases}$$

A expressão do sinal modulado obtém-se começando por considerar que o sinal modulador normalizado se escreve como uma soma de duas funções porta de largura $t_0/3$: uma positiva, de amplitude $1/2$ e atrasada de $t_0/6$, outra negativa de amplitude -1 e atrasada de $t_0/2$. Assim podemos escrever

$$\bar{m}(t) = \frac{1}{2}p\left(\frac{t - t_0/6}{t_0/3}\right) - p\left(\frac{t - t_0/2}{t_0/3}\right)$$

onde $p(t)$ é a função porta típica, centrada e de amplitude unidade. Podemos então escrever o sinal modulado

$$\begin{aligned} s(t) &= 2[1 + 0.8\bar{m}(t)] \cos(500\pi t) \\ &= 2 \left[1 + 0.4p\left(\frac{t - t_0/6}{t_0/3}\right) - 0.8p\left(\frac{t - t_0/2}{t_0/3}\right) \right] \cos(500\pi t) \end{aligned}$$

O espectro do sinal modulador $m(t)$ obtém-se a partir da expressão temporal acima sabendo que $\text{TF}[p(t)] = \text{sinc}(f)$ e portanto que devido a um atraso t_0 e a uma largura de porta τ temos que

$$\text{TF} \left[p\left(\frac{t - t_0}{\tau}\right) \right] = \tau e^{-j2\pi f t_0} \text{sinc}(f\tau)$$

de onde aplicando à expressão de $\bar{m}(t)$ acima obtemos

$$\begin{aligned} \bar{M}(f) &= \frac{t_0}{6} e^{-j\pi f t_0/3} \text{sinc}\left(\frac{t_0 f}{3}\right) - \frac{t_0}{3} e^{-j\pi f t_0} \text{sinc}\left(\frac{t_0 f}{3}\right) \\ &= \frac{t_0}{3} e^{-j\pi f t_0/3} \text{sinc}\left(\frac{t_0 f}{3}\right) [1/2 - e^{-j2\pi f t_0/3}] \end{aligned}$$

Por outro lado a expressão do espectro do sinal modulado obtem-se a partir da equação anterior deslocando-a de $+f_c$ e de $-f_c$ e adicionando os Diracs às mesmas frequências, correspondentes à portadora e tendo em conta $A_c = 2$,

$$S(f) = \delta(f - f_c) + \delta(f + f_c) + \\ 0.8 \frac{t_0}{3} e^{-j\pi(f-f_c)t_0/3} \text{sinc}\left(\frac{t_0(f-f_c)}{3}\right) [1/2 - e^{-j2\pi(f-f_c)t_0/3}] + \\ 0.8 \frac{t_0}{3} e^{-j\pi(f+f_c)t_0/3} \text{sinc}\left(\frac{t_0(f+f_c)}{3}\right) [1/2 - e^{-j2\pi(f+f_c)t_0/3}]$$

que é o espectro do sinal modulado $s(t)$.

Podemos calcular igualmente a potência do sinal modulado que se exprime utilizando (2-1.35) como

$$P_s = \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} s^2(t) dt \quad (5-1.5)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} A_c^2 [1 + a\bar{m}(t)]^2 \cos^2(2\pi ft) dt \quad (5-1.6)$$

$$= \lim_{T \rightarrow \infty} \frac{1}{T} \int_{-T/2}^{T/2} [A_c^2 \cos^2(2\pi ft) dt + A_c^2 a^2 \bar{m}^2(t) \cos^2(2\pi ft) + \\ + 2A_c^2 a \bar{m}(t) \cos(2\pi ft)] dt \quad (5-1.7)$$

$$= \frac{A_c^2}{2} + \frac{A_c^2 a^2}{2} P_{\bar{m}} + 0, \quad (5-1.8)$$

$$= \frac{A_c^2}{2} + \frac{A_c^2 a^2}{2} P_{\bar{m}} + 0, \quad (5-1.9)$$

onde se assumiu que a média do sinal modulador normalizado $\bar{m}(t)$ era igual a zero.

Exemplo: voltando ao sinal $m(t)$ do exemplo anterior, calcular as potências $P_{\bar{m}}$ e P_s .

Podemos por exemplo calcular para o sinal não normalizado $m(t)$,

$$P_m = \frac{1}{T} \int_{-T/2}^{T/2} m^2(t) dt \\ = \frac{3}{t_0} \int_0^{t_0/3} 1^2 dt + \frac{3}{t_0} \int_{t_0/3}^{2t_0/3} 2^2 dt \\ = 5,$$

de onde podemos deduzir

$$P_{\bar{m}} = \frac{P_m}{2^2} = \frac{5}{4}.$$

No que diz respeito à potência do sinal $s(t)$ podemos escrever, utilizando (5-1.9), com $A_c = 2$ e $a = 0.8$

$$P_s = \frac{A_c}{2} + \frac{A_c^2 a^2}{2} P_{\bar{m}} \\ = 2 + \frac{4(0.8)^2 5}{2 \cdot 4} \\ = 3.6.$$

5.2 Modulação de amplitude com supressão de portadora (AM-CS)

Neste caso, dito AM-CS (*Amplitude Modulation - Carrier Suppressed*) temos o mesmo tipo de modulação que no caso anterior, a diferença consistindo apenas na supressão da moduladora. Assim, o sinal modulado escreve-se

$$s(t) = m(t)c(t), \quad (5-2.1)$$

onde, como anteriormente, $m(t)$ é a mensagem ou sinal modulador e $c(t)$ é a portadora. Com $c(t) = A_c \cos(2\pi f_c t)$ encontra-se facilmente o espectro de $s(t)$ como sendo

$$S(f) = \frac{A_c}{2}M(f - f_c) + \frac{A_c}{2}M(f + f_c). \quad (5-2.2)$$

A principal vantagem deste tipo de modulação é o de a potência necessária no sinal modulado ser bastante inferior mantendo no entanto a mesma quantidade de informação transmitida. É por isso um sistema de modulação mais eficaz que o anterior. A potência em $S(f)$ é

$$P_s = \frac{A_c^2}{2}P_m, \quad (5-2.3)$$

valor inferior de $A_c/2$ à do caso anterior.

5.3 Modulação de amplitude de banda lateral única (AM-SSB)

Nos dois exemplos de modulação descritos anteriormente a informação relativa à mensagem $m(t)$ encontrava-se à frequência f_c e depois repetida em $-f_c$. Isto é redundante. Na modulação em banda lateral única (SSB - single side band) pretende-se acabar com esta redundância de forma a dividir por dois a potência necessária para transmitir a mesma quantidade de informação. A figura 5.1 mostra um sinal $m(t)$ em banda base (figura 5.1(a)) e o seu espectro modulado em AM-DSB (figura 5.1(b)). Para transmitir a informação contida na figura 5.1(a) bastaria que o sinal modulado contivesse as partes superiores e inferiores do espectro de $m(t)$ resultado da passagem do espectro da figura 5.1(b), por exemplo, por um filtro passa-alto de frequência de corte f_c (representado pela linha em pontado) resultando no espectro indicado na figura 5.1(c). Óbvio um raciocínio idêntico poderia ser feito com uma filtragem passa-baixo de frequência de corte f_c e cujo resultado seria o de preservar no sinal modulado apenas a componente inferior da mensagem $m(t)$. A modulação exemplificada no primeiro caso é chamada USSB - upper SSB e no segundo é LSSB - lower SSB. Como já tivemos ocasião de ver no capítulo 2.1.5, um sinal apenas com a

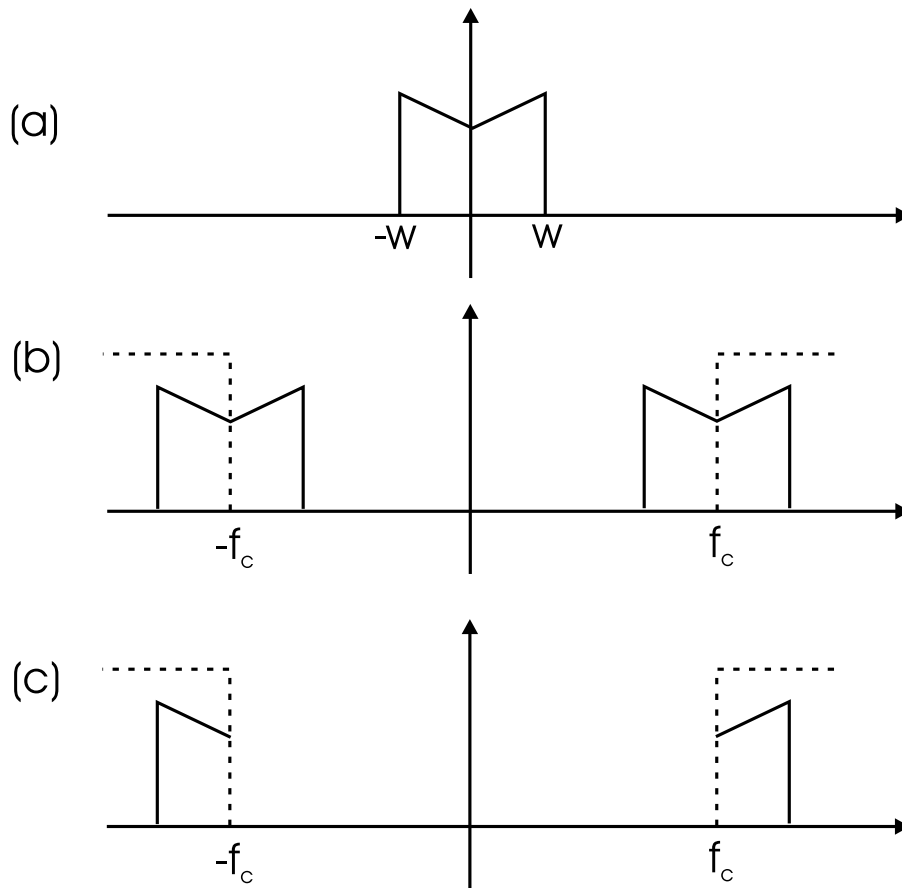


Figura 5.1: exemplificação da modulação AM-SSB.

componente espectral positiva escreve-se

$$m'(t) = \frac{1}{2}[m(t) + j\hat{m}(t)], \quad (5-3.1)$$

onde o $1/2$ serve para manter uma energia constante e $\hat{m}(t)$ é a transformada de Hilbert de $m(t)$. Neste caso o sinal modulado no domínio do tempo escreve-se

$$s'(t) = A_c m'(t) \cos(2\pi f_c t). \quad (5-3.2)$$

O espectro de $s'(t)$ obtém-se sabendo que

$$\hat{M}(f) = \begin{cases} -jM(f) & f > 0 \\ 0 & f = 0 \\ jM(f) & f < 0, \end{cases} \quad (5-3.3)$$

e portanto

$$M'(f) = \frac{1}{2}M(f) + \frac{j}{2}\hat{M}(f) \quad (5-3.4)$$

$$= \begin{cases} M(f) & f > 0 \\ 1/2M(f) & f = 0 \\ 0 & f < 0. \end{cases} \quad (5-3.5)$$

Finalmente,

$$S'(f) = \begin{cases} \frac{A_c}{2}M(f + f_c) + \frac{A_c}{2}M(f - f_c), & f_c \leq |f| \\ 0, & \text{outro } f. \end{cases} \quad (5-3.6)$$

No caso agora exemplificado apenas foram retidas as frequências positivas por isso trata-se do caso USSB. No caso LSSB teríamos

$$m'(t) = \frac{1}{2}[m(t) - j\hat{m}(t)], \quad (5-3.7)$$

e como resultado final para o espectro de $s'(t)$,

$$S'(f) = \begin{cases} \frac{A_c}{2}M(f + f_c) + \frac{A_c}{2}M(f - f_c) & f_c \geq |f| \\ 0, & \text{outro } f. \end{cases} \quad (5-3.8)$$

5.4 Modulação de amplitude em quadratura (QAM)

A última técnica de modulação em amplitude é denominada *modulação de amplitude em quadratura* (QAM - quadrature amplitude modulation). Neste caso dois sinais temporais e reais $a(t)$ e $b(t)$ formam o sinal modulador $m(t) = \frac{1}{2}[a(t) + jb(t)]$. Assim o sinal modulado escreve-se

$$s(t) = A_c m(t) \cos(2\pi f_c t) \quad (5-4.1)$$

O sinal passa banda resultante não é nem analítico nem tem simetria complexo conjugada em torno a $f = f_c$. O sinal tem, em geral, bandas laterais superior e inferior sem simetria em torno à frequência da portadora. Escreve-se neste caso a partir de (5-4.1)

$$s(t) = A_c \text{Re}\{[a(t) + jb(t)]e^{j\omega_c t}\} \quad (5-4.2)$$

$$= A_c a(t) \cos(\omega_c t) - jA_c b(t) \sin(\omega_c t) \quad (5-4.3)$$

de onde provém a denominação de modulação de amplitude em quadratura (QAM) visto tratem-se de dois sinais modulados independentemente, à mesma frequência, mas com uma defasagem relativa de $\pi/2$. Em termos de banda passante, e para a mesma banda de base, QAM requer a mesma banda que AM-DSB e o dobro da de AM-SSB, porém permite enviar dois sinais reais portadores de informação. A sua eficácia é portanto equivalente à de AM-SSB porém sem a sua dificuldade de implementação prática. QAM tornou-se por isso a técnica de modulação mais utilizada em sistemas de comunicação digital (ver capítulo 6.5).

5.5 Modulação de ângulo

Modulação de ângulo é um termo genérico para designar um tipo de modulação na qual o ângulo da portadora sinusoidal varia em função do sinal modulador, *i.e.*,

$$u(t) = A_c \cos[\phi(t)], \quad (5-5.1)$$

onde a fase $\phi(t) = f[m(t)]$, sendo $m(t)$ o sinal modulador. Dois casos são possíveis na modulação de ângulo, que são: a modulação de fase (phase modulation, PM) e a modulação de frequência (frequency modulation, FM). Em termos práticos a modulação de ângulo tem como vantagem a de a portadora manter uma amplitude constante, o que a torna mais eficiente em termos de potência transmitida no emissor, e também de ser mais imune ao ruído. É o caso bem conhecido da qualidade superior do rádio FM em relação a AM.

No caso da modulação de fase, $\phi(t)$ escreve-se

$$\phi(t) = 2\pi f_c t + \beta_a m(t), \quad (5-5.2)$$

onde β_a é a constante de modulação de fase, que se encontra ligada ao índice de modulação k_a por

$$k_a = \beta_a \max |m(t)|. \quad (5-5.3)$$

No caso da modulação de frequência, que abordaremos mais em profundidade devido à sua importância prática, $\phi(t)$ escreve-se

$$\phi(t) = 2\pi f_c t + 2\pi\beta_f \int_{-\infty}^t m(\tau) d\tau, \quad (5-5.4)$$

onde β_f é a constante de modulação e $m(t)$ é sempre o sinal modulador. Podemos assim definir a frequência instantânea

$$f_i = \frac{1}{2\pi} \frac{d\phi}{dt} = f_c + \beta_f m(t) \quad (5-5.5)$$

e em geral define-se o índice de modulação como sendo

$$k_f = \frac{\beta_f \max |m(t)|}{W}, \quad (5-5.6)$$

onde W é a largura de banda de $m(t)$. Neste caso se a mensagem $m(t)$ for tal que $|m(t)| < 1$ então podemos dizer que a constante β_f corresponde à excursão máxima em frequência em torno à portadora f_c . Isto levaria a pensar que a largura de banda ocupada pelo sinal modulado em FM seria na ordem de $2\beta_f$ o que, se fosse verdade, seria extraordinário, pois quereria dizer que a largura de banda do sinal modulado seria independente da largura de banda do sinal modulador !... O

que obviamente é impossível. Vejamos o caso particular de um sinal modulador sinusoidal $m(t) = \cos(\omega_m t)$,

$$u(t) = A_c \cos[\omega_c t + 2\pi\beta_f \int_0^t \cos(\omega\tau) d\tau] \quad (5-5.7)$$

e portanto

$$u(t) = A_c \cos[\omega_c t + \frac{2\pi\beta_f}{\omega_m} \sin(\omega_m t)] = A_c \cos[\omega_c t + \alpha \sin(\omega_m t)] \quad (5-5.8)$$

com $\alpha = \beta_f/f_m$ é o rácio entre o desvio máximo em frequência β_f e a frequência do sinal modulador. O sinal modulado pode ainda exprimir-se como

$$u(t) = \frac{A_c}{2} e^{j\omega_c t} e^{j\alpha \sin(\omega_m t)} + \frac{A_c}{2} e^{-j\omega_c t} e^{-j\alpha \sin(\omega_m t)}. \quad (5-5.9)$$

Considerando, por exemplo, apenas o termo de frequência positiva podemos dizer que, tratando-se de uma função periódica de pulsação ω_m , este se pode escrever sob forma de uma expansão em série de Fourier do tipo

$$e^{j\alpha \sin(\omega_m t)} = \sum_n C_n e^{jn\omega_m t}, \quad (5-5.10)$$

onde os coeficientes C_n dependem do coeficiente α . Esta relação entre C_n e α não é simples, mas pode-se demonstrar que se escreve como

$$C_n = J_n(\alpha), \quad (5-5.11)$$

onde $J_n(\alpha)$ é a função de Bessel de α de ordem n (ver [4]). Na prática podemos dizer que o espectro do sinal modulado se encontra centrado em $+f_c$ e $-f_c$. Em torno a cada um destes valores o espectro do sinal $m(t)$ estende-se em linhas a frequências múltiplas e sub-múltiplas de f_m com amplitudes que dependem de (5-5.11). Em princípio a decomposição em série de Fourier (5-5.10) tem um número infinito de termos. Porém na prática $|C_n|$ decresce rapidamente para $|n| > \alpha$ e considerar apenas uma banda $\pm(1 + \alpha)f_m$ é suficiente em termos práticos. Assim a banda total do sinal modulado escreve-se

$$W = 2f_m(1 + \alpha) = 2f_m + 2\alpha, \quad (5-5.12)$$

e por isso é igual a duas vezes a excursão máxima em frequência mais duas vezes a banda do sinal modulador. Apesar de não se tratar de uma demonstração rigorosa podemos deduzir que a banda útil de um sinal modulado em FM é aproximadamente

$$W = 2B + 2\alpha, \quad (5-5.13)$$

onde B é a banda do sinal modulador.

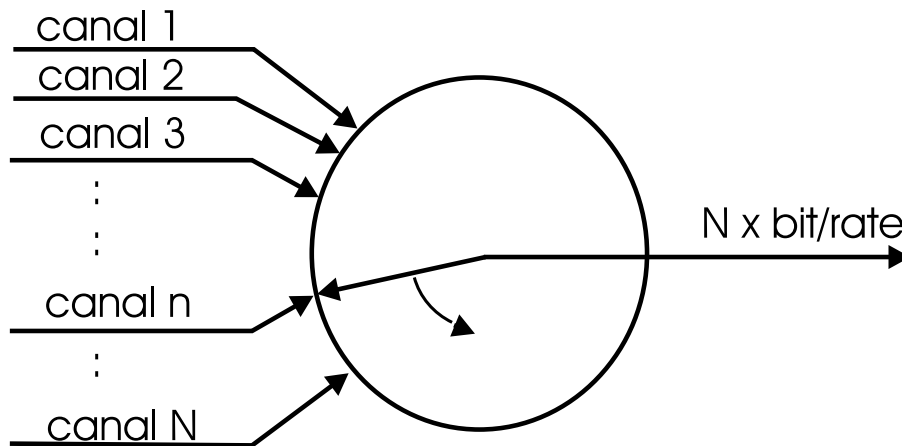


Figura 5.2: croquis de um sistema de multiplexagem temporal.

5.6 Multiplexagem no tempo e na frequência

A multiplexagem é uma técnica que permite enviar várias mensagens distintas através do mesmo canal de transmissão simultaneamente. A multiplexagem pode fazer-se seja no tempo seja na frequência. No tempo torna-se óbvio que o envio de N mensagens pelo mesmo canal requer um canal de transmissão cujo bit/rate é pelo menos $N \times$ o bit rate de cada um dos canais. Na realidade o bit rate total será sempre superior a esse número pois torna-se necessária a inclusão de bits de controlo, de correcção de erros, de identificação dos canais, etc... Um dispositivo mecânico básico que exemplifica a multiplexagem temporal encontra-se representado na figura 5.2. Na prática a multiplexagem temporal é efectuada electrónicamente por dispositivos de comutação. Mais importante, e mais utilizada, do que a multiplexagem temporal é a multiplexagem em frequência (*frequency division multiplexing*, FDM). A figura 5.3 exemplifica um tal sistema. Cada uma das mensagens é modulada em amplitude a uma frequência portadora diferente. Em seguida os sinais de cada um dos canais são somados e enviados no canal de transmissão. Torna-se óbvio que, para cada uma das mensagens possa ser recuperada intacta no receptor (desmultiplexagem), é necessário que não exista sobreposição entre espectros contíguos, i.e., que a separação entre as frequências das portadoras de cada uma das mensagens seja superior ou igual à largura de banda das mesmas mensagens. Foram definidas normas para a transmissão de canais telefónicos que agrupam linhas em bandas de frequências conforme os números de canais desejados (tabela 1).

Frequentemente utiliza-se uma técnica de modulação eficiente, tipo AM-SSB, em cada um dos canais introduzidos no FDM e depois o sinal composto obtido é de novo modulado, utilizando por exemplo uma modulação FM, antes de transmissão intercontinental através de um canal de fibra óptica ou satélite.

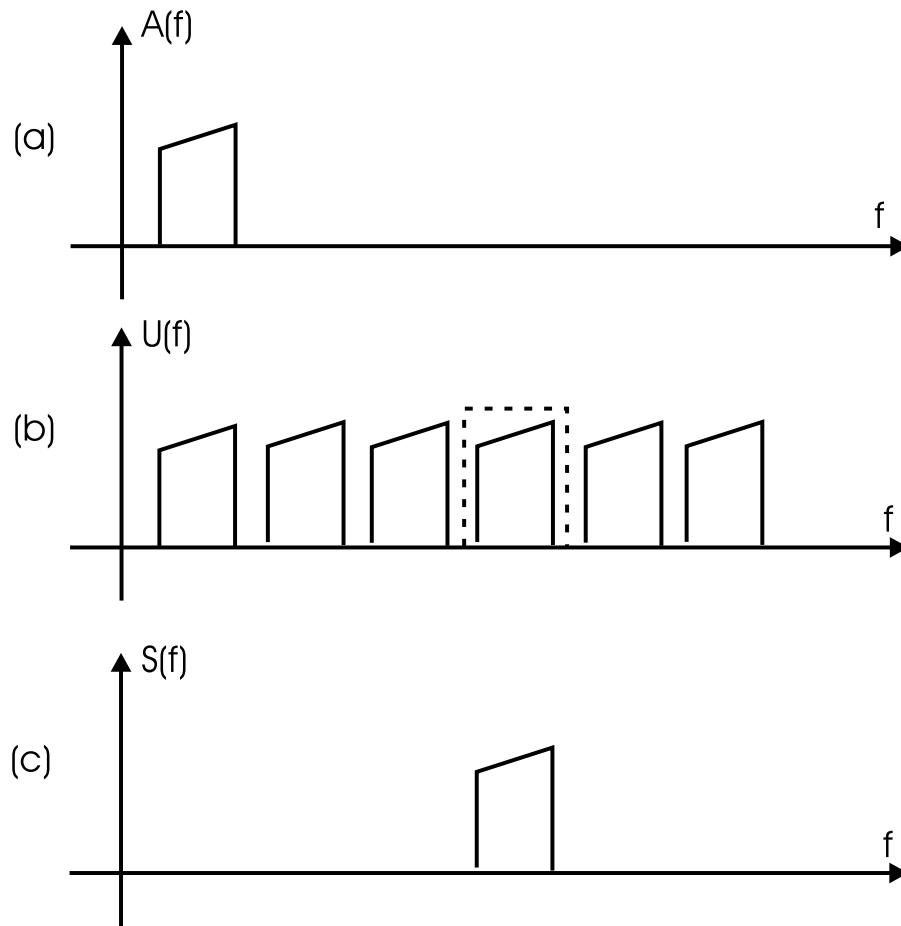


Figura 5.3: multiplexagem frequencial (FDM).

Canal		1 canal (4kHz)
Grupo		12 canais (48 kHz)
Supergrupo	5 grupos	60 canais (240kHz)
Mastergrupo	5 supergrupos	300 canais (1.2 MHz)
Supermastergrupo	3 mastergrupos	900 canais (3.6 MHz)

Tabela 1: hierarquia de FDM publicada pela CCITT.

5.7 Desmodulação AM e FM

A desmodulação de sinais modulados em AM faz-se através da extração do envelope do sinal passabanda, podendo ser efectuada de forma não coerente, i.e., sem conhecer de forma exacta a frequência do sinal modulador. Já o mesmo não acontece para a desmodulação de fase que necessita ser feita em modo coerente com o emissor, já que a informação é transmitida exactamente na fase do sinal recebido. Existe um grande

número de técnicas para efectuar a desmodulação seja de amplitude seja da fase. Na prática esta tarefa encontra-se grandemente simplificada se as componentes em fase e quadratura forem previamente extraídas utilizando o envelope complexo.

5.7.1 Extração do envelope complexo

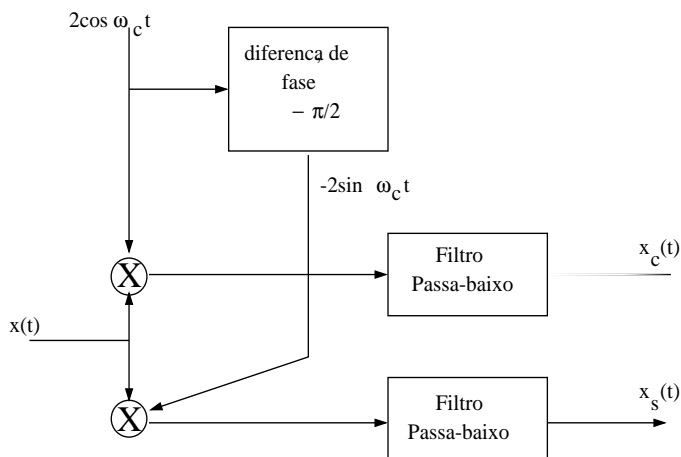


Figura 5.4: extração das componentes em fase e quadratura de um sinal passa-banda.

A importância do envelope complexo é que a partir dele podem ser obtidas todas as características do sinal passa-banda ou dos seus equivalentes passa-baixo. A extração do envelope complexo passa pelo cálculo das componentes em fase e em quadratura. A figura 5.4 ilustra esta operação. O sinal passa-banda $x(t)$ é simultaneamente multiplicado por $2\cos \omega_c t$ e por $-2\sin \omega_c t$. O resultado é uma frequência de batimento duas vezes a portadora ω_c que pode ser filtrada por um filtro passa-baixo. Com efeito o resultado do produto, por exemplo, para o $\cos \omega_c t$, é

$$\begin{aligned}
 2x(t) \cos \omega_c t &= 2 \cos \omega_c t [x_c(t) \cos \omega_c t - x_s(t) \sin \omega_c t] \\
 &= 2[x_c(t) \cos^2 \omega_c t - x_s(t) \sin \omega_c t \cos \omega_c t] \\
 &= x_c(t) + x_c(t) \cos 2\omega_c t - x_s(t) \sin 2\omega_c t, \quad (5-7.1)
 \end{aligned}$$

que é filtrado no banda útil de $x(t)$. Por exemplo se a banda de $x(t)$ em torno a f_c for $[f_c - W/2, f_c + W/2]$ então o filtro passa-baixo deverá, para retirar correctamente a componente em fase $x_c(t)$, ter uma frequência de corte f_a tal que $W/2 < f_a < 2f_c - W/2$. O mesmo raciocínio será aplicado para a componente em quadratura $x_s(t)$. O posicionamento das componentes espectrais do sinal e do batimento composto encontram-se exemplificadas na figura 5.5.

A partir das duas componentes em fase e quadratura podemos calcular o envelope

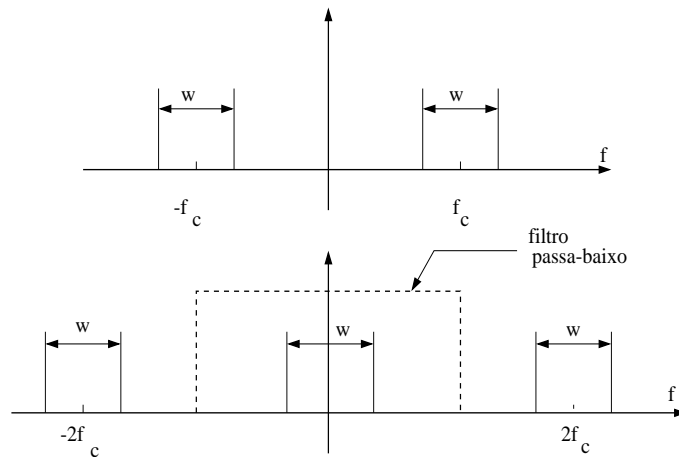


Figura 5.5: Espectros dos sinais durante a extração das componentes em fase e quadratura de um sinal passa-banda: sinal original e batimento.

complexo

$$\tilde{x}(t) = x_c(t) + jx_s(t). \quad (5-7.2)$$

5.7.2 Extração do envelope e da fase

Um detector de envelope tem a vantagem prática de não necessitar o conhecimento da fase da portadora o que simplifica enormemente o circuito necessário, que pode simplesmente ser constituído por um diodo em série com uma resistência em paralelo com um condensador. O envelope da portadora não é mais do que a amplitude da componente passa-baixo do sinal modulado, i.e., a amplitude do sinal modulador.

Qualquer sinal passa-banda se pode escrever a partir de $x_c(t)$ e $x_s(t)$, e calcular

$$a(t) = [x_c^2(t) + x_s^2(t)]^{1/2} \quad (5-7.3)$$

$$\phi(t) = \tan^{-1} \frac{x_s(t)}{x_c(t)} \quad (5-7.4)$$

onde $\phi(t)$ é a modulação de fase e a sua derivada

$$f_i(t) = \frac{d\phi(t)}{dt} \quad (5-7.5)$$

é a modulação em frequência.

Um desmodulador FM pode ser simplesmente um dispositivo que traduz a variação de frequência em variação de amplitude seguido de um detector de envelope, como explicado acima.

6 Modulação digital de sinais

6.1 Introdução

Hoje em dia a maior parte da informação encontra-se sob forma discreta, apesar de parte dessa informação ser na origem analógica ela é discretizada e quantificada na fonte. Como exemplos podemos citar a voz no caso do telefone, a música, a imagem, etc. O mais curioso é que quer a informação já discreta na origem, quer a analógica discretizada, devem ser adaptadas ao canal de transmissão que, ele, é essencialmente analógico por natureza. Este processo de adaptação ao canal de transmissão é chamado *modulação*. Já vimos exemplos de modulação analógica no capítulo anterior, e vamos agora abordar a modulação digital.

Começaremos por introduzir os conceitos preliminares de expansão de sinais em bases de funções ortogonais ajudando-nos dos conceitos análogos já apreendidos em cálculo vectorial e representações de sub-espacos vectoriais.

Passaremos em seguida em revista algumas das técnicas de base mais utilizadas hoje em dia, que são as técnicas de modulação de impulsos, quer em amplitude quer em fase, como aliás acontece no caso da modulação analógica. De uma forma geral trata-se de associar os k dígitos binários à saída do codificador e associá-los a $M = 2^k$ sinais (ou níveis) possíveis. Esta é chamada uma técnica de sinalização na qual cada um dos M níveis se encontra associado com um sinal $s_m(t)$ de um conjunto de $\{s_m(t); m = 1, \dots, M\}$ sinais possíveis. Apesar de, teoricamente, cada um dos sinais $s_m(t)$ poder ter qualquer forma, na prática, e por questões de facilidade de implementação, a família de sinais é escolhida como uma variante de uma forma de base, que é normalmente chamada “forma de impulso” ou *pulse shape*. A escolha da forma do impulso depende nalguns casos do espectro desejado para o sinal a emitir no meio físico, porque é ele que vai ser determinante na ocupação do espectro de frequências disponíveis para uma dada transmissão.

Em seguida, trataremos para cada modulação considerada, a sua representação no domínio da frequência, que é importante para determinar qual o espectro ocupado por cada uma delas, uma consideração essencial para os cada vez menos espectros disponíveis. É claro que sendo a sequência de bits a enviar de natureza aleatória, o cálculo do seu espectro torna-se num processo de estimação de espectro de potência.

6.2 Representação de sinais em expansões ortogonais

Para introduzir a expansão de sinais em bases ortogonais convém lembrar que, de forma análoga aos espaços vectoriais, nos quais podemos definir o produto vectorial e a norma, também para funções podemos escrever o produto interno entre duas

funções definidas no intervalo $[a, b]$ como sendo

$$\langle x_1(t)x_2(t) \rangle = \int_a^b x_1(t)x_2^*(t)dt, \quad (6-2.1)$$

no caso geral em que as funções podem ser complexas e onde * significa complexo conjugado. Como no caso dos vectores, se as funções forem ortogonais, temos que o seu produto interno é nulo. A partir da definição de produto interno é fácil chegar à noção de norma que é

$$\|x(t)\| = \left(\int_a^b |x(t)|^2 dt \right)^{1/2}. \quad (6-2.2)$$

Um conjunto de sinais é dito ortornormal se forem conjuntamente ortogonais e a sua norma for igual a um. Como no caso dos vectores um conjunto de sinais é dito linearmente independente, se nenhum deles pode ser representado como uma combinação linear dos outros. Adicionalmente, temos a desigualdade triangular que se escreve simplesmente

$$\|x_1(t) + x_2(t)\| \leq \|x_1(t)\| + \|x_2(t)\|, \quad (6-2.3)$$

e a desigualdade de Schwartz que se enuncia

$$\left| \int_a^b x_1(t)x_2^*(t)dt \right| \leq \left| \int_a^b |x_1(t)|^2 dt \right|^{1/2} \left| \int_a^b |x_2(t)|^2 dt \right|^{1/2}, \quad (6-2.4)$$

na qual a igualdade é obtida apenas quando $x_1(t) = ax_2(t)$, onde a é uma constante complexa.

Vamos agora generalizar ao caso de um sinal $s(t)$ real, determinístico e de energia finita \mathcal{E} ,

$$\mathcal{E} = \int_{-\infty}^{+\infty} |s(t)|^2 dt. \quad (6-2.5)$$

Podemos então aproximar o sinal $s(t)$ como uma expansão $\hat{s}(t)$ numa determinada base de funções $f_n(t)$ como

$$\hat{s}(t) = \sum_{k=1}^K s_k f_k(t) \quad (6-2.6)$$

onde a base de funções $\{f_k(t); k = 1, \dots, K\}$ é ortonormal e onde $\{s_k; k = 1, \dots, K\}$ são os coeficientes da expansão de $s(t)$ na base que se encontram através da minimização da energia do erro quadrático

$$\begin{aligned} \mathcal{E}_e &= \int_{-\infty}^{+\infty} e^2(t) dt, \\ &= \int_{-\infty}^{+\infty} [s(t) - \hat{s}(t)]^2 dt. \end{aligned} \quad (6-2.7)$$

A minimização de (6-2.7) pode ser efectuada de forma clássica diferenciando em relação a cada um dos coeficientes $\{s_k\}$ e anulando as derivadas ou, alternativamente, utilizando um resultado da teoria da estimação que consiste em considerar que o conjunto $\{s_k\}$ que minimiza a forma quadrática do erro corresponde ao caso em que a função de erro $e(t)$, é ortogonal a cada uma das funções da base $\{f_k(t)\}$. Uma explicação rápida desta noção consiste em observar que, devido às propriedades enunciadas acima, as funções $\{f_k(t)\}$ formam um sub-espço vectorial (ou um plano) no qual se encontra também $\hat{s}(t)$ visto que é uma combinação linear dos $\{f_k(t)\}$. Geometricamente a expressão (6-2.7) representa a distância entre $s(t)$ e esse espaço vectorial (esse plano), ora é sabido que a distância mínima entre um ponto e um plano é o segmento de recta que passa pelo ponto e é perpendicular ao plano. Assim, o $e(t)$ mínimo será ortogonal ao sub-espço definido pelas $\{f_k(t)\}$. Sob forma de equação podemos escrever

$$\int_{-\infty}^{+\infty} \left[s(t) - \sum_{k=1}^K s_k f_k(t) \right] f_n(t) dt = 0, \quad n = 1, \dots, K \quad (6-2.8)$$

e como, por definição as funções $\{f_k(t)\}$ são ortonormais, esta equação reduz-se a

$$s_n = \int_{-\infty}^{+\infty} s(t) f_n(t) dt, \quad n = 1, \dots, K \quad (6-2.9)$$

o que simplesmente indica que os coeficientes da expansão são obtidos através da projecção do sinal na base de funções. O erro mínimo \mathcal{E}_{\min} obtido é então dado por

$$\begin{aligned} \mathcal{E}_{\min} &= \int_{-\infty}^{+\infty} e(t) s(t) dt, \\ &= \int_{-\infty}^{+\infty} [s(t)]^2 dt - \int_{-\infty}^{+\infty} \sum_{k=1}^K s_k f_k(t) s(t) dt, \\ &= \mathcal{E}_s - \sum_{k=1}^K s_k^2. \end{aligned} \quad (6-2.10)$$

Pode-se então dizer que para $\mathcal{E}_{\min} = 0$, i.e., quando

$$\mathcal{E}_s = \sum_{k=1}^K s_k^2 = \int_{-\infty}^{+\infty} [s(t)]^2 dt, \quad (6-2.11)$$

temos que

$$s(t) = \sum_{k=1}^K s_k f_k(t), \quad (6-2.12)$$

quando o sinal $s(t)$ pode ser representado pela expressão (6-2.12) dizemos que a base de funções ortonormais $\{f_k(t)\}$ é uma **base completa**.

Eventualmente o exemplo mais conhecido de uma base de funções completa é a que resulta da expansão em série de Fourier e é um conjunto de funções trigonométricas em

seno e cosseno. O procedimento de Gram-Schmidt normalmente utilizado para vetores pode ser usado também para construir uma base completa de funções [ver disciplina de Álgebra Linear, Matemática Computacional (LESI) ou Análise Numérica (LESC)].

Vamos agora aplicar estas noções em modulação digital de sinais, onde um sinal analógico é normalmente representado na sua forma passabanda, tal que

$$s_m(t) = \text{Re}[s_{lm}(t)e^{j2\pi f_c t}], \quad (6-2.13)$$

onde $s_{lm}(t)$ é a componente passabaixa do sinal $s_m(t)$, f_c é a frequência central da modulação. Para facilitar os cálculos repetitivos dos próximos capítulos, introduzimos desde já o valor da energia \mathcal{E}_m do sinal passa banda $s_m(t)$ como

$$\begin{aligned} \mathcal{E}_m &= \int_{-\infty}^{+\infty} s_m^2(t) dt \\ &= \int_{-\infty}^{+\infty} \text{Re}[s_{lm}(t)e^{j2\pi f_c t}]^2 dt. \end{aligned} \quad (6-2.14)$$

Onde podemos agora utilizar a relação válida para qualquer número complexo z ,

$$\text{Re}[z] = \frac{1}{2}(z + z^*), \quad (6-2.15)$$

onde $*$ indica complexo conjugado e substituindo em (6-2.14)

$$\begin{aligned} \mathcal{E}_m &= \int_{-\infty}^{+\infty} \left[\frac{1}{2} [s_{lm}(t)e^{j2\pi f_c t} + s_{lm}^*(t)e^{-j2\pi f_c t}] \right]^2 dt \\ &= \frac{1}{4} \int_{-\infty}^{+\infty} [2s_{lm}(t)s_{lm}^*(t) + s_{lm}^2(t)e^{j4\pi f_c t} + [s_{lm}^*(t)]^2] e^{-j4\pi f_c t} dt, \end{aligned} \quad (6-2.16)$$

$$(6-2.17)$$

Visto que $s_{lm}(t)s_{lm}^*(t) = |s_{lm}(t)|^2$ e que $s_{lm}^2(t) = [s_{lm}^*]^2 = \text{Re}[s_m(t)]^2 + \text{Im}[s_m(t)]^2$, podemos escrever

$$\begin{aligned} \mathcal{E}_m &= \frac{1}{2} \int_{-\infty}^{+\infty} |s_{lm}(t)|^2 dt + \frac{1}{2} \int_{-\infty}^{+\infty} \{ \text{Re}[s_{lm}(t)]^2 + \text{Im}[s_{lm}(t)]^2 \} [e^{j4\pi f_c t} + e^{-j4\pi f_c t}] dt, \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} |s_{lm}(t)|^2 dt + \frac{1}{2} \int_{-\infty}^{+\infty} |s_{lm}(t)|^2 \cos[4\pi f_c t + 2\theta(t)] dt, \end{aligned} \quad (6-2.18)$$

onde, por definição, o sinal $s_{lm}(t)$ é passa-baixo comparado com a frequência f_c e assim podemos dizer que o segundo termo desta expressão é um cos que varia rapidamente, modulado lentamente em amplitude por $s_m(t)$. Baseando-nos na ideia de que o integral de um cos no intervalo $[-\infty, +\infty]$ é zero, podemos assim dizer que

o segundo termo em (6-2.17) é aproximadamente nulo o que nos dá que a energia é finalmente

$$\begin{aligned}\mathcal{E}_m &= \frac{1}{2} \int_{-\infty}^{+\infty} |s_{lm}(t)|^2 dt, \\ &= \frac{1}{2} \mathcal{E}_{lm},\end{aligned}\tag{6-2.19}$$

$$\tag{6-2.20}$$

onde \mathcal{E}_{lm} é a energia do sinal passa baixo $s_{lm}(t)$.

Tendo dois sinais funções de base $s_k(t)$ e $s_m(t)$, um dado muito útil consiste em determinar a distância Euclidiana entre eles e que é dada por

$$\begin{aligned}d_{km} &= \sqrt{\int_{-\infty}^{+\infty} [s_m(t) - s_k(t)]^2 dt} \\ &= \sqrt{\mathcal{E}_m + \mathcal{E}_k - 2\sqrt{\mathcal{E}_k \mathcal{E}_m} \text{Re}[\rho_{km}]},\end{aligned}\tag{6-2.21}$$

onde $\mathcal{E}_k = \mathcal{E}_m = \mathcal{E}$ para qualquer m e k , e ρ_{km} é o coeficiente de correlacionamento cruzado, i.e., o valor da função de correlacionamento normalizada $\phi_{mk}(\tau)$ tomada no ponto $\tau = 0$, definido por

$$\rho_{mk} = \frac{1}{2\sqrt{\mathcal{E}_k \mathcal{E}_m}} \int_{-\infty}^{\infty} s_{lm}^*(t) s_{lk}(t) dt.\tag{6-2.22}$$

A distância (6-2.21) pode ser então re-escrita

$$d_{km} = \sqrt{2\mathcal{E}[1 - \text{Re}(\rho_{km})]},\tag{6-2.23}$$

que permite ter um outro critério de diferenciação entre os dois sinais $s_m(t)$ e $s_k(t)$, para além da função de correlação, claro.

6.3 Modulação por amplitude de impulsos (*pulse amplitude modulation - PAM*)

Nesta técnica uma série de impulsos de base são modulados por uma sequência de símbolos que constituem “uma imagem” da mensagem a transmitir. Em PAM os sinais da família de impulsos são dados por

$$\begin{aligned}s_m(t) &= \text{Re}[A_m g(t) e^{j2\pi f_c t}] \\ &= A_m g(t) \cos(2\pi f_c t), \quad m = 1, \dots, M, \quad 0 \leq t \leq T_s\end{aligned}\tag{6-3.1}$$

onde $\{A_m; m = 1, \dots, M\}$ representa a sequência de bits a transmitir e que forma a amplitude de cada um dos sinais $s_m(t)$. As amplitudes A_m encontram-se distribuídas no intervalo $[(1 - M)d, (M - 1)d]$ sendo que cada nível será dado por

$$A_m = (2m - 1 - M)d, \quad m = 1, \dots, M \quad (6-3.2)$$

e portanto a diferença entre dois níveis sucessivos é de $2d$. Dado que a informação se encontra na amplitude dos impulsos, este tipo de modulação é por vezes também chamado *Amplitude Shift Keying - ASK*. O sinal $g(t)$ é a forma do impulso ou *pulse shape* da qual já falamos acima. Se a taxa de envio de bits à entrada do modulador, também chamada *bit rate*, for de R bits por segundo, o intervalo entre dois bits será de $T_b = 1/R$. À saída do modulador teremos uma taxa de símbolo (*symbol rate*) de R/k símbolos por segundo, e portanto o intervalo entre cada símbolo será $T_s = k/R = kT_b$. Começemos por definir a energia contida em cada uma das funções de base, utilizando (6-2.20)

$$\begin{aligned} \mathcal{E}_m &= \int_{-\infty}^{+\infty} s_m^2(t) dt \\ &= \frac{1}{2} \int_{-\infty}^{+\infty} |s_m(t)|^2 dt, \\ &= \frac{1}{2} \int_0^T A_m^2 g^2(t) dt, \\ &= \frac{A_m^2}{2} \int_0^T g^2(t) dt, \\ &= \frac{A_m^2 \mathcal{E}_g}{2}, \end{aligned} \quad (6-3.3)$$

onde \mathcal{E}_g é a energia do sinal $g(t)$. Para simplificar os cálculos que seguem é conveniente exprimir os impulsos de base como

$$s_m(t) = A_m f(t), \quad (6-3.4)$$

onde $f(t)$ é uma função normalizada em energia

$$f(t) = \sqrt{\frac{2}{\mathcal{E}_g}} g(t) \cos(2\pi f_c t), \quad (6-3.5)$$

e assim temos que a amplitude $s_m = A_m \sqrt{\mathcal{E}_g/2}$. Claramente a possibilidade de confundir o nível m com o nível n depende da “distância” entre $s_m(t)$ e $s_n(t)$. Dado que cada um dos impulsos só difere pela sua amplitude a “distância” euclidiana mede-se ao longo da recta, o que nos permite escrever

$$\begin{aligned} d_{mn} &= \sqrt{(s_m - s_n)^2} \\ &= |A_m - A_n| \sqrt{\mathcal{E}_g/2}. \end{aligned} \quad (6-3.6)$$

Visto que dois níveis sucessivos se encontram separados por $2d$, a distância mínima entre os sinais correspondentes é

$$d_{min} = d\sqrt{2\mathcal{E}_g}. \quad (6-3.7)$$

Na prática o sinal de base da eq. (6-3.1) é um tipo de modulação AM-DSB, no qual duas bandas (cuja forma depende de $g(t)$) em torno a $-f_c$ e $+f_c$ são ocupadas pelo espectro de $s_m(t)$. Poderíamos efectuar uma modulação PAM-SSB, da mesma forma que no caso AM-SSB, utilizando o sinal analítico ou o seu conjugado consoante desejarmos a banda lateral única superior ou inferior, respectivamente. Nesse caso teremos

$$s_m(t) = \text{Re}\{A_m[g(t) \pm j\hat{g}(t)]e^{j2\pi f_c t}\}, \quad m = 1, \dots, M, \quad 0 \leq t \leq T. \quad (6-3.8)$$

Se, por exemplo, o canal de transmissão for tal que deixe passar as frequências em torno a zero Hz não necessitamos de introduzir o termo exponencial em torno a f_c e o impulso de base escreve-se então simplesmente

$$s_m(t) = A_m g(t), \quad (6-3.9)$$

que é chamado PAM em banda de base (*baseband PAM*).

6.4 Modulação por fase de impulsos

(*pulse phase modulation - PPM*)

Neste caso e, como o seu nome indica, será a fase e não a amplitude do impulso que variará de acordo com a sequência digital a transmitir. Assim o impulso de base escreve-se

$$\begin{aligned} s_m(t) &= \text{Re}[g(t)e^{j2\pi(m-1)/M}e^{j2\pi f_c t}], \\ &= g(t) \cos[2\pi f_c t + \frac{2\pi}{M}(m-1)], \\ &= g(t) \cos \frac{2\pi}{M}(m-1) \cos 2\pi f_c t - g(t) \sin \frac{2\pi}{M}(m-1) \sin 2\pi f_c t. \end{aligned} \quad (6-4.1)$$

Assim a fase θ_m da portadora pode tomar os M valores, $\theta_m = 2\pi(m-1)/M$; $m = 1, \dots, M$. Normalmente, PPM é uma modulação também chamada *Phase-Shift Keying* (PSK), que adoptaremos no que segue. Podemos notar, fazendo um cálculo análogo ao efectuado no caso da PAM, que cada um dos impulsos utilizados em PSK têm um energia constante

$$\begin{aligned} \mathcal{E} &= \int_0^T s_m^2(t) dt, \\ &= \frac{1}{2} \int_0^T g^2(t) dt, \\ &= \frac{\mathcal{E}_g}{2}. \end{aligned} \quad (6-4.2)$$

Utilizando a última forma de (6-4.1) podemos notar que cada função de base pode ser representada como uma combinação linear de duas funções ortogonais $f_1(t)$ e $f_2(t)$, tais que

$$s_m(t) = s_{m1}f_1(t) + s_{m2}f_2(t), \quad (6-4.3)$$

onde

$$f_1(t) = \sqrt{\frac{2}{\mathcal{E}_g}}g(t) \cos(2\pi f_c t), \quad (6-4.4)$$

$$f_2(t) = -\sqrt{\frac{2}{\mathcal{E}_g}}g(t) \sin(2\pi f_c t), \quad (6-4.5)$$

e finalmente, tomando como coordenadas as funções $f_1(t)$ e $f_2(t)$ podemos representar o sinal $s_m(t)$ como um vector \mathbf{s}_m de coordenadas

$$\begin{aligned} \mathbf{s}_m &= \left[\sqrt{\frac{\mathcal{E}_g}{2}} \cos \frac{2\pi}{M}(m-1), \sqrt{\frac{\mathcal{E}_g}{2}} \sin \frac{2\pi}{M}(m-1) \right] \\ &= [s_{m1}, s_{m2}], \end{aligned} \quad (6-4.6)$$

Torna-se assim fácil compreender a representação dos impulsos de base num referencial ortonormado de acordo com a figura 6.1 para $M=2, 4$ e 8 . Note-se que os símbolos atribuídos a cada ponto do espaço encontram-se distribuídos de tal forma a só mudar um bit entre pontos contíguos, com a intenção de minimizar o erro cometido. Da

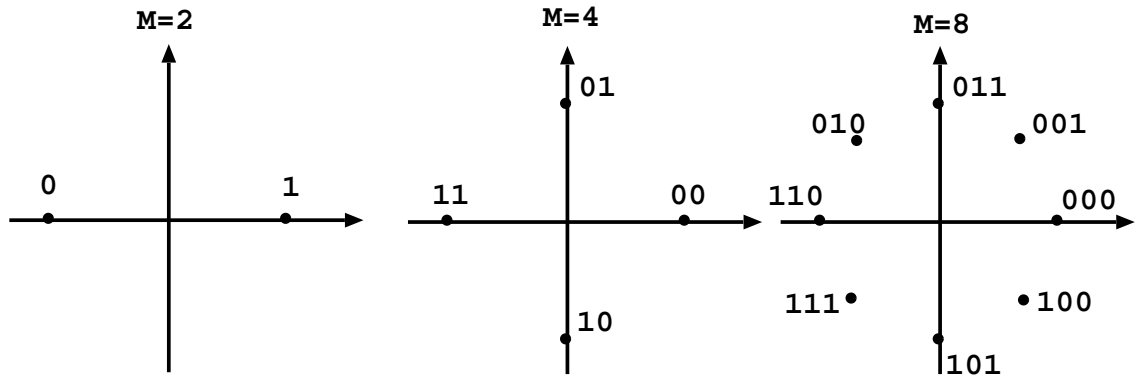


Figura 6.1: diagrama de espaço de sinais para a modulação PSK.

mesma forma que com PAM, em PSK, o erro entre dois níveis sucessivos depende da distância euclidiana entre dois pontos no plano, dada por

$$\begin{aligned} d_{mn} &= |s_m - s_n| \\ &= \left\{ \mathcal{E}_g \left[1 - \cos \frac{2\pi}{M}(m-n) \right] \right\}^{1/2} \end{aligned} \quad (6-4.7)$$

onde o seu valor mínimo é obtido para $m = n + 1$ de onde

$$d_{min} = \sqrt{\mathcal{E} \left(1 - \cos \frac{2\pi}{M}\right)}. \quad (6-4.8)$$

6.5 Modulação de amplitude em quadratura (*Quadrature Amplitude Modulation - QAM*)

A modulação de amplitude em quadratura (QAM) permite obter a mesma eficiência da modulação PAM-SSB utilizando dois bits de informação em duas portadoras em quadratura (90 graus) uma da outra. É um pouco como se fosse PAM em PSK com $M = 2$. Neste caso os sinais de base escrevem-se

$$\begin{aligned} s_m(t) &= \operatorname{Re}[(A_{mc} + jA_{ms})g(t)e^{j2\pi f_c t}], \quad m = 1, \dots, M, \quad 0 \leq t \leq T, \\ &= A_{mc}g(t) \cos(2\pi f_c t) - A_{ms}g(t) \sin(2\pi f_c t). \end{aligned} \quad (6-5.1)$$

Deste modo cada impulso leva o dobro da informação através da codificação de A_{mc} e A_{ms} , amplitudes das componentes em fase e em quadratura da portadora, respectivamente. Outra forma de ver a relação entre QAM, PAM e PSK é de escrever o sinal da equação (6-5.1) como

$$\begin{aligned} s_m(t) &= \operatorname{Re}[V_m g(t) e^{j\theta_m} e^{j2\pi f_c t}], \\ &= V_m g(t) \cos(2\pi f_c t + \theta_m), \end{aligned} \quad (6-5.2)$$

onde $V_m = \sqrt{A_{mc}^2 + A_{ms}^2}$ e $\theta_m = \arctan(A_{ms}/A_{mc})$, o que demonstra que QAM pode ser visto como uma combinação de modulação de amplitude e modulação de fase.

Podemos agora generalizar a ideia introduzida com QAM a ordens superiores de PAM e PSK, combinando sempre os dois tipos de modulação. Assim, se notarmos M_1 o número de níveis de PAM, e M_2 o número de níveis de PSK, temos que o número de níveis total obtido pela combinação PAM-PSK, é $M = M_1 M_2$. Isto significa que se podem representar $k = m_1 + m_2$ símbolos onde $M_1 = 2^{m_1}$ e $M_2 = 2^{m_2}$. A taxa de transmissão de símbolos encontra-se então reduzida a $1/T_s = R/(m_1 + m_2)$. A título de exemplo, a figura 6.2 representa os diagramas de espaço de sinais para a modulação combinada PAM-PSK, para $M = 8$ e $M = 16$. Como no caso PSK, também em QAM os sinais de base $s_m(t)$ podem ser representados em forma vectorial a partir de um conjunto das funções de base (6-4.5) com as coordenadas

$$\mathbf{s}_m = [A_{mc}\sqrt{\mathcal{E}/2}, A_{ms}\sqrt{\mathcal{E}/2}] \quad (6-5.3)$$

onde, como habitualmente \mathcal{E}_g é a energia do impulso de base $g(t)$. A distância euclidiana entre dois pontos m, n é neste caso dada por

$$\begin{aligned} d_{mn} &= |\mathbf{s}_m - \mathbf{s}_n|, \\ &= \sqrt{\frac{\mathcal{E}}{2} [(A_{mc} - A_{nc})^2 + (A_{ms} - A_{ns})^2]}, \end{aligned} \quad (6-5.4)$$

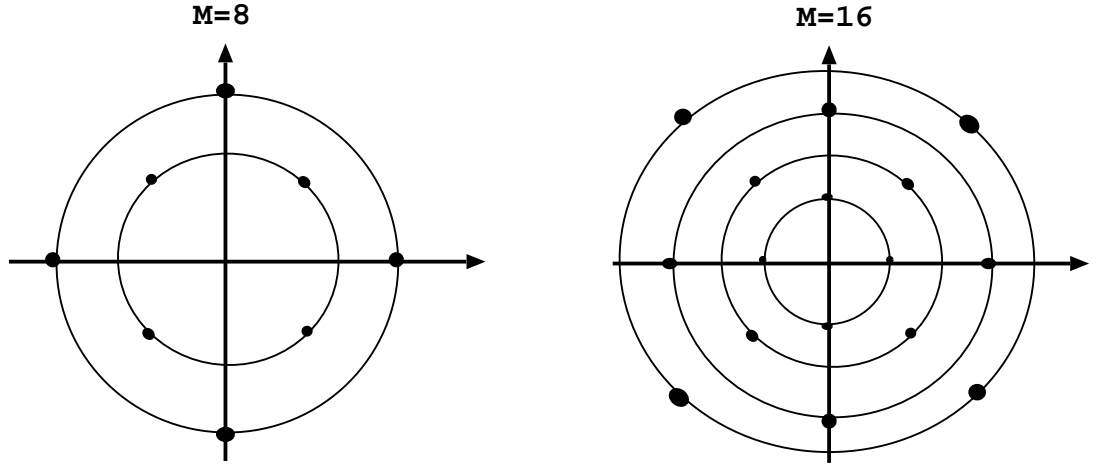


Figura 6.2: diagrama de espaço de sinais para a modulação PAM-PSK, $M=8$ e $M=16$.

A distância mínima entre dois pontos contíguos no espaço de sinais QAM depende da forma da disposição da grelha de pontos. No caso em que as amplitudes discretas variam de acordo com $\{(2m - 1 - M)d; m = 1, \dots, M\}$ a grelha toma uma forma rectangular e nesse caso a distância mínima é dada por

$$d_{min} = d\sqrt{2\mathcal{E}_g}. \quad (6-5.5)$$

6.6 Modulação de frequência (*Frequency Shift Keying - FSK*)

Neste último caso que consideraremos aqui, a separação entre níveis faz-se através de variações em frequência. Assim, a família de sinais de base escreve-se

$$\begin{aligned} s_m(t) &= \text{Re}[s_{lm}(t)e^{j2\pi f_c t}], \quad m = 1, \dots, M, \quad 0 \leq t \leq T, \\ &= \sqrt{\frac{2\mathcal{E}_g}{T}} \cos(2\pi f_c t + 2\pi m \delta f t), \end{aligned} \quad (6-6.1)$$

onde as formas de onda passa-baixo toma a forma

$$s_{lm}(t) = \sqrt{\frac{2\mathcal{E}_g}{T}} e^{j2\pi m \Delta f t}, \quad (6-6.2)$$

Para o caso em que $\Delta f = 1/2T$ os M sinais FSK podem ser representados como outros tantos vectores de dimensão M , dados por

$$\mathbf{s}_1 = [\sqrt{\mathcal{E}}, 0, 0, \dots, 0]$$

$$\begin{aligned} \mathbf{s}_2 &= [0, \sqrt{\mathcal{E}}, 0, \dots, 0] \\ &\vdots \\ \mathbf{s}_M &= [0, 0, 0, \dots, 0, \sqrt{\mathcal{E}}], \end{aligned} \tag{6-6.3}$$

$$\tag{6-6.4}$$

A distância no espaço de sinais é dada por

$$d_{mn} = \sqrt{2\mathcal{E}}, \tag{6-6.5}$$

e que sendo independente de m e n é também a distância mínima.

6.7 Densidade espectral de potência do sinal modulado

Começemos por notar que podemos designar um sinal de tipo passabanda modulado em torno à frequência f_c como

$$s(t) = \text{Re}[s_l(t)e^{j2\pi f_c t}], \tag{6-7.1}$$

onde $s_l(t)$ é um sinal passabaixo de $s(t)$, onde para as modulações lineares introduzidas acima se escreve

$$s_l(t) = \sum_{m=-\infty}^{\infty} A_m g(t - mT_s). \tag{6-7.2}$$

De notar que para este tipo de modulações A_m é real e contém a informação da mensagem no caso PAM, e é complexo no caso PSK, QAM, e PAM-PSK. Em qualquer dos casos a sequência $\{A_m, m = 1, \dots, M\}$ é aleatória e portanto $s_l(t)$ e $s(t)$ são eles mesmos processos aleatórios. O primeiro passo no cálculo da densidade espectral de potência é estimar a função de autocorrelação, que para o caso do sinal (6-7.1) se escreve

$$\phi_s(\tau) = \text{Re}[\phi_{sl}(\tau)e^{j2\pi f_c \tau}] \tag{6-7.3}$$

onde $\phi_{sl}(\tau)$ é a função de autocorrelação do componente passa-baixo $s_l(t)$. A TF de (6-7.3) permite obter a densidade espectral de potência

$$\begin{aligned} P_s(f) &= \text{TF}[\phi_s(\tau)], \\ &= \frac{1}{2}[P_{sl}(f - f_c) + P_{sl}(f + f_c)], \end{aligned} \tag{6-7.4}$$

onde a se utilizou explicitamente a expressão (6-2.15). Trata-se agora de determinar a expressão da densidade espectral de potência da componente passa-baixo $P_{sl}(f)$. Para começar notamos que

$$E[s_l(t)] = \sum_{m=-\infty}^{\infty} E[A_m]g(t - mT) = 0, \tag{6-7.5}$$

visto que $E[A_m] = 0$. Agora a função de autocorrelação de $s_l(t)$ escreve-se

$$\phi_{sl}(t + \tau, t) = E[s_l(t + \tau)s_l^*(t)], \quad (6-7.6)$$

e por substituição de (6-7.2) em (6-7.6) obtemos

$$\phi_{sl}(t + \tau, t) = \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \phi_A(m - n)g(t + \tau - mT)g^*(t - nT), \quad (6-7.7)$$

onde

$$\phi_A(m - n) = E[A_m A_n^*], \quad (6-7.8)$$

fazendo uma primeira mudança de variável $i = m - n$, obtemos

$$\phi_{sl}(t + \tau, t) = \sum_{i=-\infty}^{\infty} \phi_A(i) \sum_{n=-\infty}^{\infty} g(t + \tau - iT - nT)g^*(t - nT), \quad (6-7.9)$$

e agora uma segunda mudança de variável $u = t - nT$,

$$\phi_{sl}(u + nT + \tau, u + nT) = \sum_{i=-\infty}^{\infty} \phi_A(i) \sum_{n=-\infty}^{\infty} g(u + \tau - iT)g^*(u), \quad (6-7.10)$$

e podemos facilmente deduzir que ϕ_{sl} é uma função periódica de período T , o que significa que $s(t)$ é um processo periodicamente estacionário ou cicloestacionário. Como tivemos ocasião de referir anteriormente, o cálculo do espectro dos sinais cicloestacionários faz-se através da TF da média da função de autocorrelação do processo num período. Essa média escreve-se neste caso como

$$\bar{\phi}_{sl}(\tau) = \frac{1}{T} \int_0^T \phi_{sl}(t + \tau, t) dt, \quad (6-7.11)$$

e por isso substituindo (6-7.10) em (6-7.11) com a mudança de variável $t = u + nT$, o que implica que quando $t = 0$, $u = -nT$ e para $t = T$, $u = -(n - 1)T$ e então

$$\bar{\phi}_{sl}(\tau) = \sum_{i=-\infty}^{\infty} \phi_A(i) \sum_{n=-\infty}^{\infty} \frac{1}{T} \int_{-nT}^{-(n-1)T} g(u + \tau - iT)g^*(u) du, \quad (6-7.12)$$

de onde, combinando o somatório em n e o integral num intervalo T , podemos escrever

$$\bar{\phi}_{sl}(\tau) = \sum_{i=-\infty}^{\infty} \phi_A(i) \int_{-\infty}^{\infty} \frac{1}{T} g(u + \tau - iT)g^*(u) du, \quad (6-7.13)$$

e finalmente

$$\bar{\phi}_{sl}(\tau) = \frac{1}{T} \sum_{i=-\infty}^{\infty} \phi_A(i) \phi_g(\tau - iT), \quad (6-7.14)$$

onde

$$\phi_g(\tau) = \int_{-\infty}^{\infty} g(t + \tau)g^*(t) dt. \quad (6-7.15)$$

Escrevendo finalmente o espectro de potência de $s_l(t)$ como a TF de $\bar{\phi}_{sl}(\tau)$, *i.e.*,

$$P_{sl}(f) = \int_{-\infty}^{\infty} \bar{\phi}_{sl}(\tau) e^{-j2\pi f\tau} d\tau, \quad (6-7.16)$$

e notando que (6-7.14) representa apenas a convolução das autocorrelações de A_k e da função de pulso $g(t)$, permite chegar a

$$P_{sl}(f) = \frac{1}{T} |G(f)|^2 P_A(f), \quad (6-7.17)$$

e que para este caso, em virtude de $P_A(f) = \sigma_A^2/T$, temos

$$P_{sl}(f) = \frac{\sigma_A^2}{T} |G(f)|^2, \quad (6-7.18)$$

de onde podemos concluir que o espectro do sinal a transmitir é perfeitamente independente da mensagem dependendo apenas do espectro da função de pulso $g(t)$. Na prática somos levados a determinar $g(t)$ de forma a que o seu espectro se adapte ao canal de transmissão.

6.8 Exemplos

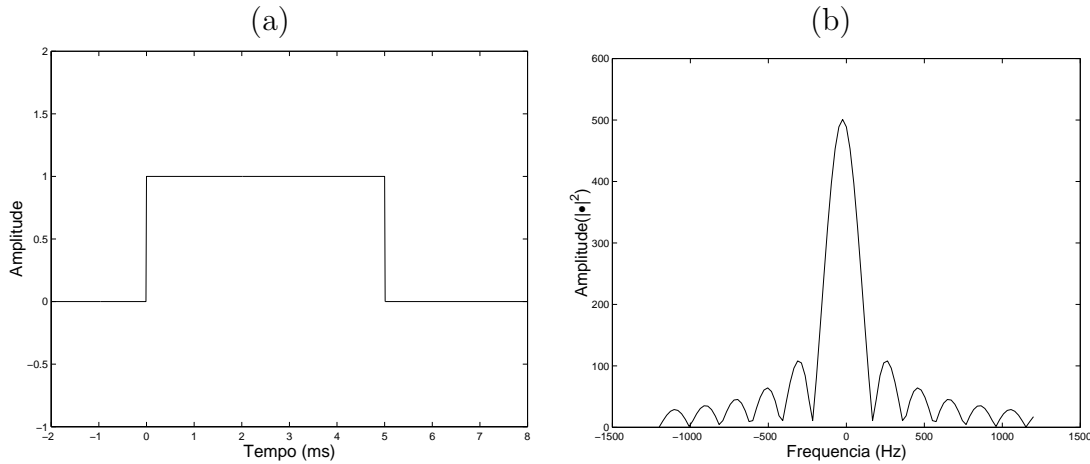


Figura 6.3: forma de pulso rectangular causal de duração $T = 5$ ms (a) e respectivo espectro de potência $|G(f)|^2$ (b).

Exemplo 1: considere a forma de pulso representada na figura 6.3(a), e definida por

$$g(t) = \begin{cases} 0 & t \in [-\infty, 0[\\ A & t \in [0, T] \\ 0 & t \in]T, \infty[\end{cases} \quad (6-8.1)$$

Calcule o seu espectro de potência $|G(f)|^2$.

Trata-se de um cálculo clássico, que se pode fazer como exercício de revisão da TF, assim temos que

$$\begin{aligned}
 G(f) &= \int_0^T A e^{-j2\pi ft} dt, \\
 &= A \left[\frac{e^{-j2\pi ft}}{-j2\pi f} \right]_0^T \\
 &= A \left[\frac{e^{-j2\pi fT} - 1}{-j2\pi f} \right] \\
 &= A e^{-j\pi fT} \left[\frac{e^{-j\pi fT} - e^{j\pi fT}}{-j2\pi f} \right] \\
 &= A e^{-j\pi fT} \frac{\sin(\pi fT)}{\pi f}, \tag{6-8.2}
 \end{aligned}$$

$$\tag{6-8.3}$$

ou ainda

$$G(f) = AT \frac{\sin \pi fT}{\pi fT} e^{-j\pi fT} \tag{6-8.4}$$

e finalmente

$$|G(f)| = (AT)^2 \left(\frac{\sin \pi fT}{\pi fT} \right)^2 \tag{6-8.5}$$

espectro que se encontra representado na figure 6.3(b).

Exemplo 2: outra forma de pulso tipicamente utilizada em comunicações é o *raised cosine* dado pela expressão

$$g(t) = \begin{cases} \frac{A}{2} \left[1 + \cos \frac{2\pi}{T} \left(t - \frac{T}{2} \right) \right], & 0 \leq t \leq T \\ 0 & t \notin [0, T], \end{cases} \tag{6-8.6}$$

e representada na figura 6.4(a). O cálculo do seu espectro de potência passa obviamente pela TF de $g(t)$. O cálculo é algo semelhante ao do exemplo anterior porém muito mais trabalhoso. Após algumas páginas de cálculo podemos chegar ao resultado notável

$$G(f) = \frac{AT}{2} \frac{\sin \pi fT}{\pi fT(1 - f^2T^2)} e^{-j\pi fT}, \tag{6-8.7}$$

sendo o seu espectro de potência

$$G(f) = \left(\frac{AT}{2} \right)^2 \left(\frac{\sin \pi fT}{\pi fT(1 - f^2T^2)} \right)^2, \tag{6-8.8}$$

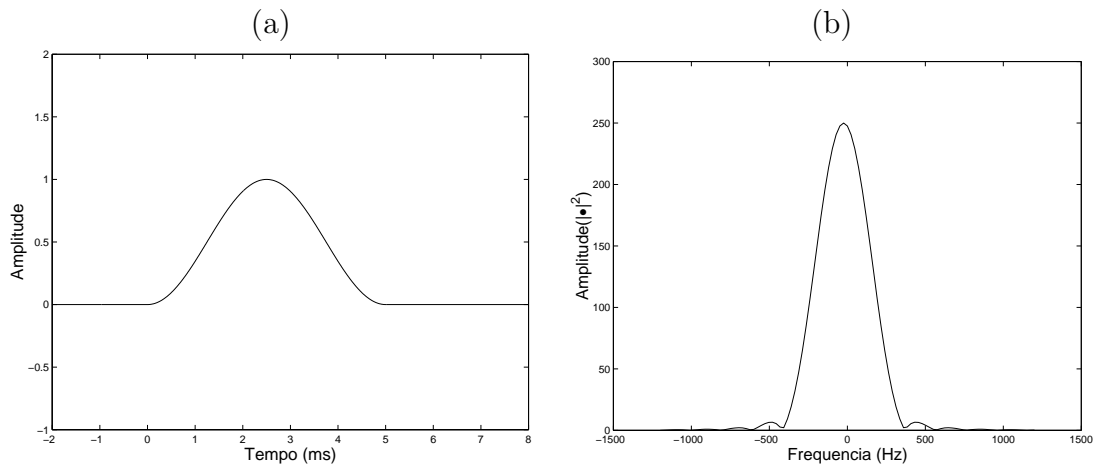


Figura 6.4: forma de pulso *raised cosine* de duração $T = 5$ ms (a) e respectivo espectro de potência $|G(f)|^2$ (b).

que se encontra representado na figura 6.4(b). Comparando com o espectro de potência da forma de pulso rectangular pode-se notar que também aqui os nulos são obtidos para múltiplos do inverso da duração do pulso n/T , porém contrariamente ao caso do pulso rectangular o primeiro zero não se obtém para $n = 1$ mas só a partir de $n = 2$, iniciando uma largura do lóbulo principal muito maior do que no caso do pulso rectangular. Em compensação, a curva de decaimento do lóbulo central do espectro de potência é muito mais rápida do que no caso do pulso rectangular. Em comunicações este facto é apreciável dado que permite obter uma menor interferência entre pulsos.

7 Receptor óptimo em presença de ruído aditivo

Como já tivemos oportunidade de referir o desmodulador no bloco receptor, constitui o elemento central do sistema de comunicação pois ele permite realizar, para além da colocação em banda base através de uma deslocação em frequência do sinal saído do canal de transmissão, duas tarefas essenciais:

- eliminar ou pelo menos diminuir a ISI à entrada do amostrador e do decisor trabalhando em conjunto com o filtro emissor e a definição da função de pulso.
- aumentar a relação sinal ruído à entrada do decisor e assim diminuir a probabilidade da escolha de um símbolo errado.

Já vimos igualmente que estas duas tarefas, apesar de não serem completamente contraditórias, necessitam o estabelecimento de um compromisso tendo em conta a complexidade prática dos sistema a implementar. O problema da equalização do canal de transmissão para a diminuição da ISI vai ser deixado para a cadeira de Comunicações Digitais. Aqui vamos apenas abordar a questão da maximização da probabilidade de detecção à entrada do decisor assumindo que não existe ISI e que o canal é do tipo ruído aditivo Gaussiano e, nesse caso, vamos determinar qual o receptor óptimo. Veremos então que esse receptor óptimo pode ser implementado através do receptor correlador ou filtro adaptado (*matched filter*).

7.1 O receptor óptimo

Como já foi dito acima, supor que o sistema não tem ISI é equivalente a supor que cada pulso (por exemplo num sistema PAM) é recebido isoladamente e não tem relação com o seguintes, numa palavra, trata-se de um sistema sem memória. De um modo geral o que se pretende é que dado um determinado sinal recebido

$$r(t) = s_m(t) + w(t), \quad (7-1.1)$$

cuja versão discreta num determinado intervalo de tempo pode ser escrita sob forma vectorial $\mathbf{r} = \mathbf{s}_m + \mathbf{w}$, a probabilidade de que seja tomada a decisão de que o sinal $s_m(t)$ se encontra presente seja tomada. Vamos notar essa probabilidade como $P[\mathbf{s}_m|\mathbf{r}]$, que por palavras é a probabilidade para que o sinal \mathbf{s}_m seja escolhido quando o sinal \mathbf{r} é observado sabendo que \mathbf{s}_m foi emitido. Esta probabilidade é normalmente designada por probabilidade à posteriori. O critério que consiste em maximizar a probabilidade à posteriori (e que veremos corresponde a minimizar a probabilidade de erro) é denominado o *maximum a posteriori probability (MAP) criterium*. Utilizando

a lei de Bayes podemos escrever

$$P[\mathbf{s}_m|\mathbf{r}] = \frac{P[\mathbf{r}|\mathbf{s}_m]P[\mathbf{s}_m]}{P[\mathbf{r}]}, \quad (7-1.2)$$

onde $P[\mathbf{r}|\mathbf{s}_m]$ é a probabilidade de receber a sequência \mathbf{r} condicionada no sinal \mathbf{s}_m . Por oposição $P[\mathbf{s}_m]$ é normalmente denominada probabilidade à priori. A resolução do problema requer o conhecimento da probabilidade à priori $P[\mathbf{s}_m]$. No caso de sinais equiprováveis (frequente na prática), onde $P[\mathbf{s}_m] = 1/M; m = 1, \dots, M$, substituindo em (7-1.2) permite obter

$$P[\mathbf{s}_m|\mathbf{r}] = \frac{1}{M} \frac{P[\mathbf{r}|\mathbf{s}_m]}{P[\mathbf{r}]}, \quad (7-1.3)$$

de onde se vê que o problema de encontrar \mathbf{s}_m que a maximiza a probabilidade à posteriori é idêntico ao de maximizar a densidade de probabilidade condicional $P[\mathbf{r}|\mathbf{s}_m]$, dado que $P[\mathbf{r}]$ não depende de \mathbf{s}_m . É interessante verificar que o critério MAP conduz à maximização da mesma densidade de probabilidade do máximo de verosimilhança, neste sentido $P[\mathbf{r}|\mathbf{s}_m]$ é denominada função de verosimilhança (ou *likelihood function*). A maximização da função de verosimilhança conduz ao critério de máxima verosimilhança (*maximum likelihood (ML) criterion*). Neste caso, no qual os sinais \mathbf{s}_m são equiprováveis, os critérios MAP e ML conduzem ao mesmo resultado. Considerando que o ruído aditivo introduzido pelo canal segue uma distribuição gaussiana, podemos notar que cada amostra r_k do sinal $r(t)$ tem média s_{mk} dado que o ruído é de média nula. Assim

$$P[r_k] = \frac{1}{\sqrt{\pi\sigma^2}} \exp \left[-\frac{(r_k - s_{mk})^2}{\sigma^2} \right], \quad m = 1, \dots, M \quad (7-1.4)$$

onde σ^2 é a variância do ruído e, por consequente, a variância de r_k dado que s_{mk} é assumido determinístico. Notando que \mathbf{r} é formado por N amostras não correlacionadas, podemos escrever a densidade de probabilidade conjunta do vector \mathbf{r} como o produto das densidades marginais,

$$\begin{aligned} P[\mathbf{r}|\mathbf{s}_m] &= \prod_{k=1}^N P[r_k] \\ &= \frac{1}{(\pi\sigma^2)^{N/2}} \exp \left[-\sum_{k=1}^N \frac{(r_k - s_{mk})^2}{\sigma^2} \right], \quad m = 1, \dots, M \end{aligned} \quad (7-1.5)$$

A partir deste ponto torna-se conveniente escrever a função de verosimilhança utilizando uma função monótona da densidade de probabilidade como por exemplo o logaritmo de (7-1.5) (frequentemente chamada *log-likelihood function*). A partir da equação precedente temos que o log da função de verosimilhança se pode escrever

$$\begin{aligned} \log P[\mathbf{r}|\mathbf{s}_m] &= -\frac{N}{2}(\pi\sigma^2) - \frac{1}{\sigma^2} \sum_{k=1}^N (r_k - s_{mk})^2, \quad m = 1, \dots, M \\ &= -\frac{N}{2}(\pi\sigma^2) - \frac{1}{\sigma^2} \|\mathbf{r} - \mathbf{s}_m\|^2, \quad m = 1, \dots, M \end{aligned} \quad (7-1.6)$$

onde se nota que o primeiro termo é constante e independente de \mathbf{s}_m e onde a maximização da expressão é equivalente à minimização do segundo termo (por causa do sinal - antes do somatório) em relação a \mathbf{s}_m . Este segundo termo escreve-se (deprezando a constante multiplicativa),

$$\begin{aligned} C(\mathbf{r}, \mathbf{s}_m) &= \|\mathbf{r} - \mathbf{s}_m\|^2 \quad m = 1, \dots, M \\ &= |\mathbf{r}|^2 - 2\mathbf{r} \cdot \mathbf{s}_m + |\mathbf{s}_m|^2, \quad m = 1, \dots, M \end{aligned} \quad (7-1.7)$$

sendo que o primeiro termo da expressão à direita não depende de \mathbf{s}_m podemos notar que a minimização de $C(\mathbf{r}, \mathbf{s}_m)$ é equivalente à maximização de $C'(\mathbf{r}, \mathbf{s}_m)$ dada por

$$C'(\mathbf{r}, \mathbf{s}_m) = 2\mathbf{r} \cdot \mathbf{s}_m - |\mathbf{s}_m|^2, \quad m = 1, \dots, M \quad (7-1.8)$$

onde $\mathbf{r} \cdot \mathbf{s}_m; m = 1, \dots, M$ é o produto escalar entre os dois vectores \mathbf{r} e \mathbf{s}_m para todos os valores de m ou, por outras palavras, a projecção de \mathbf{r} no sub espaço vectorial gerado pelos vectores $\mathbf{s}_m; m = 1, \dots, M$ e onde o segundo termo $|\mathbf{s}_m|^2$ representa a energia contida em cada um dos sinais $s_m(t)$. O valor de cada uma destas projecções pode ser visto como a correlação entre o sinal recebido $r(t)$ e cada um dos possíveis sinais emitidos $s_m(t)$, e assim C' pode ser dado por um (chamado) receptor-correlacionador

$$C'(\mathbf{r}, \mathbf{s}_m) = 2 \int_0^T r(t)s_m(t)dt - \mathcal{E}_m \quad m = 1, \dots, M \quad (7-1.9)$$

onde a função de correlação é calculada para o ponto $\tau = 0$ e em seguida é seleccionado o sinal s_m para o qual $C(\mathbf{r}, \mathbf{s}_m)$ é máximo. O receptor óptimo correspondente encontra-se representado na figura 7.1.

7.2 O filtro adaptado como receptor óptimo

A maximização da relação sinal/ruído à saída de um filtro é um problema clássico em teoria do sinal e pode ser transposto para o caso da recepção de mensagens num sistema de comunicação considerando primeiramente que: o sistema se encontra em banda base (ou seu equivalente) e que apenas um símbolo é transmitido, o que significa que não existe ISI. O sinal recebido pode ser então

$$y(t) = A_0 h(t) + n(t), \quad (7-2.1)$$

onde A_0 é o símbolo transmitido, $h(t)$ é a forma de pulso depois de passada no canal e $n(t)$ é o ruído suposto aditivo branco e Gaussiano. Neste caso o sinal à saída do filtro de recepção de resposta impulsiva $f(t)$ escreve-se

$$q(t) = \int_{-\infty}^{\infty} y(\tau)f(t - \tau)d\tau. \quad (7-2.2)$$

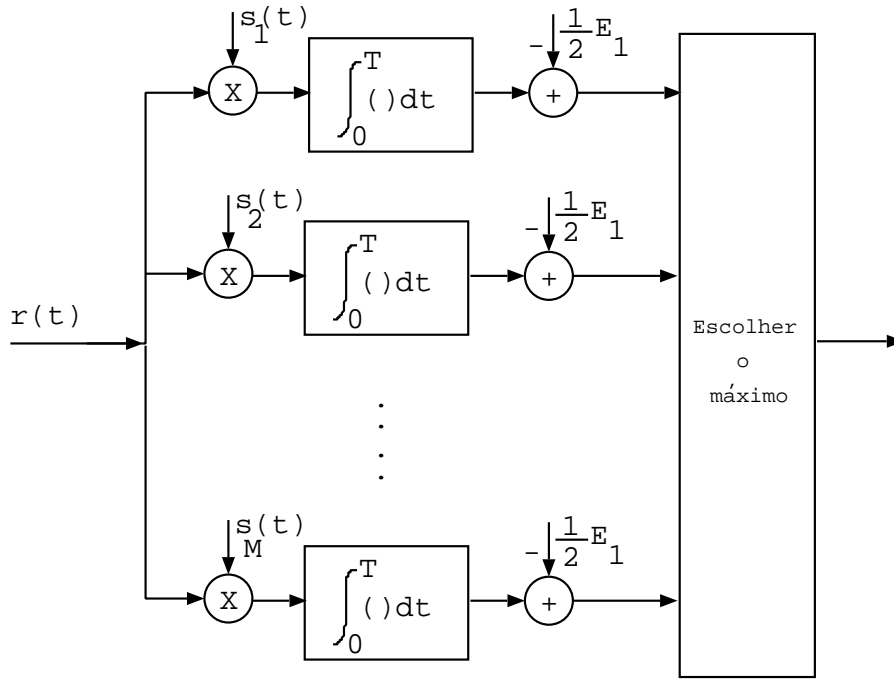


Figura 7.1: receptor correlador para M sinais.

O sinal $q(t)$ encontra-se depois amostrado ao instante $t = 0$ (para retirar o símbolo A_0) e então

$$Q_0 = \int_{-\infty}^{\infty} y(\tau)f(-\tau)d\tau. \quad (7-2.3)$$

Substituindo agora (7-2.1) em (7-2.3) temos

$$Q_0 = A_0 \int_{-\infty}^{\infty} h(\tau)f(-\tau)d\tau + \int_{-\infty}^{\infty} n(\tau)f(-\tau)d\tau, \quad (7-2.4)$$

na qual o primeiro termo corresponde ao sinal e o segundo corresponde ao ruído. Intuitivamente o desempenho do nosso decisor será tanto melhor quanto maior for o primeiro termo em relação ao segundo, *i.e.*, quanto maior for a relação SNR à saída do filtro. Esta relação SNR escreve-se

$$\text{SNR}_{\text{out}} = \frac{\text{energia sinal}}{\text{energia ruído}}, \quad (7-2.5)$$

onde ambas as energias são medidas à saída do filtro. A energia do sinal E_s é dada pela $E[\cdot]$ do primeiro termo de (7-2.4), *i.e.*,

$$E_s = \sigma_A^2 \left| \int_{-\infty}^{\infty} h(\tau)f(-\tau)d\tau \right|^2, \quad (7-2.6)$$

enquanto a energia do ruído (suposto estacionário e ergódico), E_n , à saída do filtro é dada por

$$E_n = E[n_0^2] = \langle n_0^2 \rangle = \int_{-\infty}^{\infty} |F(f)|^2 P_n(f) df, \quad (7-2.7)$$

onde $\langle . \rangle$ indica média temporal e onde o ruído branco $P_n(f) = \text{cte} = N_0$ e por isso, utilizando de novo o teorema de Parseval

$$E_n = N_0 \int_{-\infty}^{\infty} f^2(\tau) d\tau. \quad (7-2.8)$$

O problema agora é de determinar de forma única o filtro $f(t)$ que maximize

$$\text{SNR}_{\text{out}} = \frac{\sigma_A^2 \left| \int_{-\infty}^{\infty} h(\tau) f(-\tau) d\tau \right|^2}{N_0 \int_{-\infty}^{\infty} f^2(\tau) d\tau}. \quad (7-2.9)$$

A demonstração mais simples passa pela utilização da desigualdade de Schwarz que diz que

$$\left| \int_a^b g_1(x) g_2(x) dx \right|^2 \leq \left[\int_a^b |g_1(x)|^2 dx \right] \left[\int_a^b |g_2(x)|^2 dx \right], \quad (7-2.10)$$

onde a igualdade só é verificada quando $g_2(x) = K g_1(x)$, K constante. Portanto, utilizando (7-2.10) no numerador de (7-2.9) chegamos à conclusão que o valor máximo do SNR é obtido quando

$$f(t) = K h(-t). \quad (7-2.11)$$

Este valor do filtro $f(t)$ é chamado filtro adaptado ou *matched filter*. Mais, quando $f(t)$ é o *matched filter* então obtemos o valor máximo do SNR que é

$$\text{SNR}_{\text{max}} = \frac{\sigma_A^2 \sigma_h^2}{N_0}, \quad (7-2.12)$$

onde σ_h^2 é a energia na função de pulso recebida

$$\sigma_h^2 = \int_{-\infty}^{\infty} |h(t)|^2 dt. \quad (7-2.13)$$

O *matched filter* tem como resposta impulsiva a forma do sinal de entrada revertida no tempo (time-reversal), $h(-t)$ e a sua TF é $H^*(\omega)$ isto é o conjugado da forma do sinal de entrada. Podemos assim perceber que a resposta do filtro adaptado "equaliza" exactamente a fase do sinal de entrada resultando numa saída que é real e igual a $|H(\omega)|^2$. O valor do sinal de saída do filtro adaptado escreve-se a partir de (7-2.3) como

$$Q_0 = \int_{-\infty}^{\infty} y(\tau) h(\tau) d\tau, \quad (7-2.14)$$

o que não é mais do que o valor da função de correlação entre a entrada e a forma de pulso esperada para o instante $t = 0$. Por isso o filtro adaptado é por vezes também chamado receptor correlador.

Demonstrámos que o filtro adaptado constitui o receptor óptimo de um único impulso em ruído branco e Gaussiano. Este resultado pode ser facilmente generalizado ao caso de uma sucessão de impulsos desde que não exista ISI. Para que tal não aconteça é necessário e suficiente que cada impulso $h(t)$ do sinal recebido esteja contido num e só num intervalo de símbolo.

7.3 Em banda passante

No caso do sinal PAM em banda passante um único pulso recebido no receptor escreve-se

$$y(t) = \sqrt{2}\text{Re}\{A_0h(t)e^{j\omega_c t}\} + n(t), \quad (7-3.1)$$

onde agora o pulso recebido $h(t)$ e o símbolo A_0 podem ser complexos. Demonstra-se que o correlador-receptor ou matched filter maximiza o SNR à saída do filtro receptor e segue exactamente os mesmos passos que no caso banda base mas onde os $[\cdot]^2$ foram alterados em $|\cdot|^2$ devido ao facto de $h(t)$ ser agora uma função complexa.

Desafiamos o leitor a fazer essa demonstração que resulta no evidente resultado que o matched filter se escreve neste caso

$$f(t) = h^*(-t). \quad (7-3.2)$$

A Tabelas e relações particulares

A.1 Transformada de Fourier

- 67 -
Table III : Paires de Transformées de Fourier

	Fonctions du Temps, $f(t)$		Fonctions de Fréquence, $F(\omega)$	
1	$\frac{1}{\tau} p_{\tau/2}(t) = \frac{1}{\tau} [u_1(t + \frac{\tau}{2}) - u_1(t - \frac{\tau}{2})]$ $= \frac{1}{\tau} \text{rect}(\frac{t}{\tau})$			$\frac{\sin(\frac{\tau\omega}{2})}{(\frac{\tau\omega}{2})} = \frac{\sin \tau\pi f}{\tau\pi f}$ $= \text{sinc}(\tau f)$
2	$\frac{1}{\tau} \text{sinc}(\frac{t}{\tau}) = \frac{1}{\tau} \frac{\sin(\pi \frac{t}{\tau})}{(\pi \frac{t}{\tau})}$			$p_{\tau/\pi}(\omega) = \text{rect}(\frac{\omega}{2\pi/\tau})$ $= \text{rect}(\tau f)$
3	$q_{\tau}(t) = \begin{cases} \frac{1}{\tau} - \frac{ t }{\tau^2}, & t < \tau \\ 0, & t > \tau \end{cases}$			$\frac{\text{sinc}^2(\frac{\tau\omega}{2})}{(\frac{\tau\omega}{2})^2}$
4	$\frac{1}{\tau} \frac{\text{sinc}^2(\frac{\pi t}{\tau})}{(\frac{\pi t}{\tau})^2}$			$q_{\tau/\pi}(\omega) = \begin{cases} 1 - \frac{ \omega }{2\pi/\tau}, & \omega < \frac{2\pi}{\tau} \\ 0, & \omega > \frac{2\pi}{\tau} \end{cases}$
5	$\frac{1}{\tau\sqrt{2\pi}} e^{-\frac{1}{2}(\frac{t}{\tau})^2}$			$e^{-\frac{1}{2}(\tau\omega)^2}$
6	$\frac{2}{\pi\tau} \sqrt{1 - (\frac{t}{\tau})^2}, t < \tau$ $0, t > \tau$			$\frac{2 J_1(\omega\tau)}{(\omega\tau)}$
7	$\alpha e^{-\alpha t} u_1(t)$			$\frac{\alpha}{\alpha + j\omega}$
8	$\frac{\alpha}{2} e^{-\alpha t }$			$\frac{\alpha^2}{\alpha^2 + \omega^2}$
9	$\frac{(\alpha/\pi)}{1 + \alpha^2 t^2}$			$e^{-\frac{ \omega }{\alpha}}$
10	$\alpha^2 t e^{-\alpha t} u_1(t)$			$\frac{\alpha^2}{(\alpha + j\omega)^2}$
11	$e^{-\alpha t} \sin \beta t \cdot u_1(t)$			$\frac{\beta}{(\alpha^2 + \beta^2) - \omega^2 + j 2\alpha\omega}$
12	$e^{-\alpha t} [\cos \beta t - \frac{\alpha}{\beta} \sin \beta t] u_1(t)$			$\frac{j\omega}{(\alpha^2 + \beta^2) - \omega^2 + j 2\alpha\omega}$

Table III (Suite)

	Fonctions du Temps, $f(t)$	Fonctions de Fréquence, $F(\omega)$		
13.	$\text{sgn } t = \frac{t}{ t }$			$\frac{2}{j\omega}$
14.	$\frac{1}{jt}$			$-\pi \text{sgn } \omega$
15.	$u_1(t) = \frac{1}{2} + \frac{1}{2} \text{sgn } t$			$\pi u_0(\omega) + \frac{1}{j\omega}$
16.	$u_0(t)$ $u_0^{(n)}(t) \triangleq \frac{d^n u_0(t)}{dt^n}$			$\frac{1}{(j\omega)^n}$
17.	$\frac{t}{t^n}$			$\frac{2\pi u_0(\omega)}{2\pi j^n u_0^{(n)}(\omega)}$
18.	$u_1(t) * \frac{1}{T} P_{T/2}(t)$			$\pi u_0(\omega) + \frac{2 \sin(\omega T/2)}{j T \omega^2}$
19.	$\cos \omega_0 t$			$\pi [u_0(\omega + \omega_0) + u_0(\omega - \omega_0)]$
20.	$\sin \omega_0 t$			$j\pi [u_0(\omega + \omega_0) - u_0(\omega - \omega_0)]$
21.	$[\cos \omega_0 t] \cdot u_1(t)$			$\frac{\pi}{2} [u_0(\omega + \omega_0) + u_0(\omega - \omega_0)] + \frac{j\omega}{\omega_0^2 - \omega^2}$
22.	$[\sin \omega_0 t] \cdot u_1(t)$			$j\frac{\pi}{2} [u_0(\omega + \omega_0) - u_0(\omega - \omega_0)] + \frac{\omega_0}{\omega_0^2 - \omega^2}$
23.	$\sum_{-\infty}^{\infty} u_0(t - nT)$			$\omega_0 \sum_{-\infty}^{\infty} u_0(\omega - n\omega_0)$ ($\omega_0 T = 2\pi$)
24.	Fonction Périodique $f(t) = f(t + T) = \sum_{-\infty}^{\infty} \alpha_n e^{jn\omega_0 t}$ $\alpha_n = \frac{1}{T} \int_0^T f(t) e^{-jn\omega_0 t} dt$			$2\pi \sum_{-\infty}^{\infty} \alpha_n u_0(\omega - n\omega_0)$

A.2 Relações Trigonométricas Usuais

$$\begin{aligned} \sin^2 x + \cos^2 x &= 1 & \cos^2 x &= \frac{1 + \cos 2x}{2} & \sin^2 x &= \frac{1 - \cos 2x}{2} \\ \sin x &= \frac{e^{jx} - e^{-jx}}{2j} & \cos x &= \frac{e^{jx} + e^{-jx}}{2} & \tan x &= \frac{e^{jx} - e^{-jx}}{j(e^{jx} + e^{-jx})} \end{aligned}$$

Adição

$$\begin{aligned} \sin(a + b) &= \sin a \cos b + \sin b \cos a & \sin(a - b) &= \sin a \cos b - \sin b \cos a \\ \cos(a + b) &= \cos a \cos b - \sin a \sin b & \cos(a - b) &= \cos a \cos b + \sin a \sin b \\ \tan(a + b) &= \frac{\tan a + \tan b}{1 - \tan a \tan b} & \tan(a - b) &= \frac{\tan a - \tan b}{1 + \tan a \tan b} \end{aligned}$$

Multiplicação: com $\tan a = t$

$$\sin(2a) = 2 \sin a \cos a = \frac{2t}{1 + t^2}$$

$$\cos(2a) = \cos^2 a - \sin^2 a = 2 \cos^2 a - 1 = 1 - 2 \sin^2 a = \frac{1 - t^2}{1 + t^2}$$

$$\tan(2a) = \frac{2 \tan a}{1 - \tan^2 a} = \frac{2t}{1 - t^2}$$

$$\cos a \cos b = \frac{1}{2} [\cos(a + b) + \cos(a - b)] \quad \sin a \sin b = \frac{1}{2} [\cos(a - b) - \cos(a + b)]$$

$$\sin a \cos b = \frac{1}{2} [\sin(a + b) + \sin(a - b)]$$

$$\cos p + \cos q = 2 \cos \frac{p + q}{2} \cos \frac{p - q}{2} \quad \cos p - \cos q = -2 \sin \frac{p + q}{2} \sin \frac{p - q}{2}$$

$$\sin p + \sin q = 2 \sin \frac{p + q}{2} \cos \frac{p - q}{2} \quad \sin p - \sin q = 2 \sin \frac{p - q}{2} \cos \frac{p + q}{2}$$

$$\tan p + \tan q = \frac{\sin(p + q)}{\cos p \cos q} \quad \tan p - \tan q = \frac{\sin(p - q)}{\cos p \cos q}$$

Trigonometria Hiperbólica

$$\cosh x + \sinh x = \exp(x) \quad \cosh x - \sinh x = \exp(-x) \quad \cosh^2 x - \sinh^2 x = 1$$

$$\begin{aligned}
\sinh(a+b) &= \sinh a \cosh b + \sinh b \cosh a & \sinh(a-b) &= \sinh a \cosh b - \sinh b \cosh a \\
\cosh(a+b) &= \cosh a \cosh b + \sinh a \sinh b & \cosh(a-b) &= \cosh a \cosh b - \sinh a \sinh b \\
\sinh 2a &= 2 \sinh a \cosh a & \cosh 2a &= \cosh^2 a - \sinh^2 a = 1 + 2 \sinh^2 a = 2 \cosh^2 a - 1 \\
\cosh^2 a &= \frac{1 + \cosh 2a}{2} & \sinh^2 a &= \frac{\cosh 2a - 1}{2} & \tanh^2 a &= \frac{\cosh 2a - 1}{\cosh 2a + 1} \\
& & \tanh 2a &= \frac{2 \tanh a}{1 + \tanh^2 a} \\
\sinh x &= \frac{e^x - e^{-x}}{2} & \cosh x &= \frac{e^x + e^{-x}}{2} & \tanh x &= \frac{e^x - e^{-x}}{e^x + e^{-x}} \\
\cosh jx &= \cos x & \sinh jx &= j \sin x & \tanh jx &= j \tan x
\end{aligned}$$

A.3 Desenvolvimentos em série

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1} \epsilon(x)$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \dots + (-1)^n \frac{x^{2n}}{(2n)!} + x^{2n} \epsilon(x)$$

$$\tan x = x - \frac{x^3}{3} + \frac{2x^5}{15} + x^5 \epsilon(x)$$

$$\sinh x = x + \frac{x^3}{3!} + \dots + \frac{x^{2n-1}}{(2n-1)!} + x^{2n-1} \epsilon(x)$$

$$\cosh x = 1 + \frac{x^2}{2!} + \dots + \frac{x^{2n}}{(2n)!} + x^{2n} \epsilon(x)$$

$$\tanh x = x - \frac{x^3}{3} + \frac{2x^5}{15} + x^5 \epsilon(x)$$

$$\arcsin x = x + \frac{x^3}{6} + \dots + \frac{1.3.5 \dots (2n+1)}{2.4.6 \dots (2n)} \frac{x^{2n+1}}{(2n+1)!} - x^{2n+1} \epsilon(x)$$

$$\arctan x = x - \frac{x^3}{3} + \frac{x^5}{5} - \dots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1} \epsilon(x)$$

$$\sinh^{-1} x = x - \frac{x^3}{6} + \dots + (-1)^n \frac{1.3.5 \dots (2n+1)}{2.4.6 \dots (2n)} \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1} \epsilon(x)$$

$$\tanh^{-1} x = x + \frac{x^3}{3} + \frac{x^5}{5} + \dots + \frac{x^{2n+1}}{(2n+1)!} + x^{2n+1} \epsilon(x)$$

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} + x^n \epsilon(x)$$

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \dots + (-1)^{n-1} \frac{x^n}{n} + x^n \epsilon(x)$$

$$(1+x)^\alpha = 1 + \alpha x + \alpha(\alpha-1) \frac{x^2}{2} + \dots + \alpha(\alpha-1) \dots (\alpha-n+1) \frac{x^n}{n} + x^n \epsilon(x)$$

$$\frac{1}{1-x} = 1 + x + x^2 + \dots + x^n + x^n \epsilon(x)$$

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-1)^n x^n + x^n \epsilon(x)$$

A.4 Algumas relações úteis

A.4.1 Integrais

$$\int_0^\infty e^{-ax^2} dx = \frac{1}{2} \sqrt{\frac{\pi}{a}}$$

$$\int_0^\infty x e^{-ax^2} dx = \frac{1}{2a}$$

$$\int_0^\infty x^2 e^{-ax^2} dx = \frac{\sqrt{\pi}}{4a^{3/2}}$$

$$\int_0^\infty x^3 e^{-ax^2} dx = \frac{1}{2a^2}$$

$$\int_0^\infty x^4 e^{-ax^2} dx = \frac{3}{8a^2} \sqrt{\frac{\pi}{a}}$$

A.4.2 Séries

Geométrica: $u_1 + qu_1 + q^2u_1 + \dots + q^{n-1}u_1 = u_1 \frac{1-q^n}{1-q}$

Aritmética: $u_1 + qu_1 + 2qu_1 + \dots + (n-1)qu_1 =$

A.4.3 Derivadas

$$[a^{f(x)}]' = \log a a^{f(x)} f'(x)$$

A.4.4 Trigonometria do círculo

	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin x$	0	1/2	$\sqrt{2}/2$	$\sqrt{3}/2$	1
$\cos x$	1	$\sqrt{3}/2$	$\sqrt{2}/2$	1/2	0
$\tan x$	0	$\sqrt{3}/3$	1	$\sqrt{3}$	∞
$\cot x$	∞	$\sqrt{3}$	1	$\sqrt{3}/3$	0

B Séries de Fourier

Um sinal periódico é definido pela relação intuitiva

$$\exists T_0 < \infty \quad \text{tal que} \quad s(t) = s(t + nT_0), \quad (\text{B-0.1})$$

onde T_0 é chamado período do sinal e n é um número inteiro.

Pode-se demonstrar que todas as funções periódicas que

1. sejam uniformes
2. sejam finitas ou com discontinuidades absolutamente integráveis, i.e., que

$$\int_t^{t+T_0} |s(u)| du < \infty$$

3. tenham um número finito de discontinuidades
4. tenham um número finito de máximos e mínimos num período

podem ser representadas por uma combinação linear arbitrária de funções sinusoidais de frequências crescentes. Assim, todo sinal periódico pode ser representado pela série de Fourier

$$s(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos n\omega_0 t + b_n \sin n\omega_0 t), \quad (\text{B-0.2})$$

onde

$$a_n = \frac{2}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} s(t) \cos n\omega_0 t dt, \quad (\text{B-0.3})$$

e

$$b_n = \frac{2}{T_0} \int_{-\frac{T_0}{2}}^{\frac{T_0}{2}} s(t) \sin n\omega_0 t dt, \quad (\text{B-0.4})$$

onde T_0 é o período do sinal $s(t)$ e ω_0 é a pulsação definida por $\omega_0 = 2\pi/T_0$. Outra forma equivalente do desenvolvimento em série de Fourier de $s(t)$ é

$$s(t) = \sum_{n=0}^{\infty} A_n \cos(n\omega_0 t + \phi_n). \quad (\text{B-0.5})$$

Exercício: utilizando B-0.3 e B-0.4 demonstrar que na relação B-0.5 temos

$$A_0 = \frac{a_0}{2},$$

$$A_n = \sqrt{a_n^2 + b_n^2},$$

$$\phi_n = \arctan\left(-\frac{b_n}{a_n}\right).$$

Finalmente a terceira forma, chamada forma complexa, da série de Fourier escreve-se

$$s(t) = \sum_{n=-\infty}^{+\infty} C_n e^{jn\omega_0 t}, \quad (\text{B-0.6})$$

com

$$C_n = \frac{1}{T_0} \int_{-\frac{T_0}{2}}^{+\frac{T_0}{2}} s(t) e^{-jn\omega_0 t} dt. \quad (\text{B-0.7})$$

Exercício: demonstrar a relação B-0.7.

Exemplo 1: calcular o desenvolvimento em série de Fourier da onda quadrada, periódica, de valor médio nulo, de período T_0 e amplitude E representada na figura B.1.

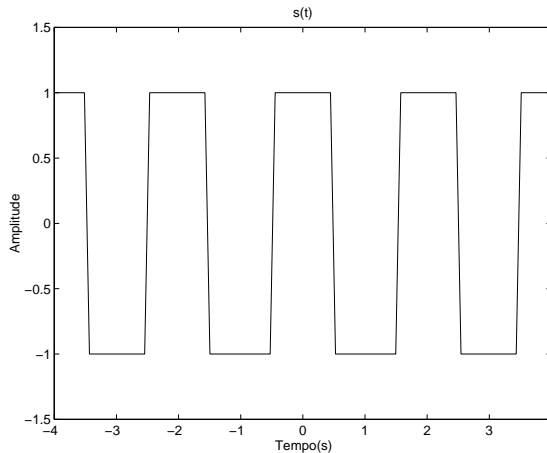


Figura B.1: onda quadrada.

Observando o sinal da figura B.1 podemos desde já determinar que $a_0 = 0$ pois este coeficiente é proporcional ao valor médio do sinal que neste caso é nulo. Por outro lado, visto que o sinal representado é uma função par, temos que o desenvolvimento em série deverá ser também uma função par e que portanto $b_n = 0$. Calculemos então a_n . Utilizando B-0.3

$$a_n = -\frac{2}{T_0} \int_{-T_0/2}^{-T_0/4} \cos n\omega_0 t dt + \frac{2}{T_0} \int_{-T_0/4}^{T_0/4} \cos n\omega_0 t dt - \frac{2}{T_0} \int_{T_0/4}^{T_0/2} \cos n\omega_0 t dt,$$

com $T_0 = 2$. Integrando o coseno nos devidos intervalos e pondo em factor $1/n\omega_0$ obtem-se

$$a_n = -\frac{1}{n\pi}[-\sin(n\pi/2) + \sin(n\pi) - \sin(n\pi/2) - \sin(n\pi/2) + \sin(n\pi) - \sin(n\pi/2)],$$

sabendo que $\sin n\pi = 0$ e juntando os termos em $\sin(n\pi/2)$

$$a_n = \frac{4}{n\pi} \sin \frac{n\pi}{2}.$$

O desenvolvimento de $s(t)$ escreve-se então

$$s(t) = \frac{4}{\pi} \left(\cos \omega_0 t - \frac{1}{3} \cos 3\omega_0 t + \frac{1}{5} \cos 5\omega_0 t + \dots + \frac{1}{n} \sin n \frac{\pi}{2} \cos n\omega_0 t + \dots \right),$$

com $\omega_0 = 2\pi/T_0 = \pi$.

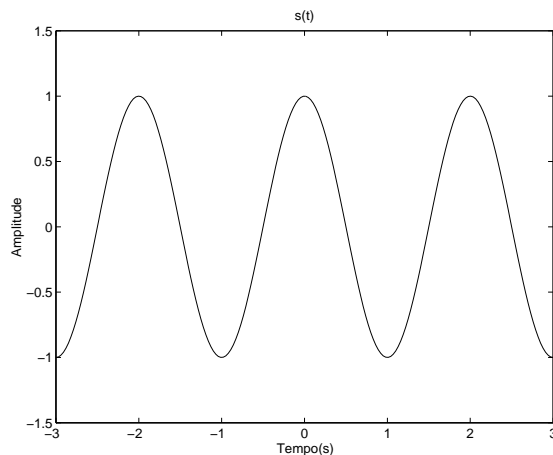


Figura B.2: sinal sinusoidal.

Exemplo 2: calcular o desenvolvimento em série de Fourier da onda sinusoidal, de valor médio nulo, periódica, de período T_0 e amplitude unidade representada na figura B.2. Como no exemplo anterior, trata-se de um sinal par, de valor médio nulo e portanto $a_0 = b_n = 0$. Assim

$$a_n = \frac{2}{T_0} \int_{-T_0/2}^{T_0/2} \cos \omega_0 t \cos n\omega_0 t dt,$$

com $T_0 = 2$. Sabendo que $\cos a \cos b = 1/2[\cos(a+b) + \cos(a-b)]$ (ver anexo) obtemos facilmente, após integração,

$$a_n = \frac{\sin(n+1)\pi}{(n+1)\pi} + \frac{\sin(n-1)\pi}{(n-1)\pi},$$

ou numa forma mais compacta

$$a_n = \begin{cases} 1 & n = \pm 1 \\ 0 & n \neq \pm 1 \end{cases}$$

Utilizando B-0.2 obtemos (como aliás seria evidente desde o início)

$$s(t) = \cos \omega_0 t.$$

O coeficiente C_n da equação B-0.7 é uma função de ordem n da variável discreta $n\omega_0$ (harmónica n de $s(t)$) e dá uma representação (discreta) do sinal $s(t)$ no domínio da frequência. Por essa razão é geralmente chamado espectro do sinal periódico $s(t)$. C_n é em geral uma função complexa e contém, nesse caso, a informação de amplitude e de fase das componentes (harmónicas) de $s(t)$ que se pode escrever sob a forma

$$C_n(n\omega_0) = \|C_n(n\omega_0)\| e^{j\phi_n(n\omega_0)}, \quad (\text{B-0.8})$$

onde $\|C_n(n\omega_0)\|$ é o espectro de amplitude e $\phi_n(n\omega_0)$ é o espectro de fase de $s(t)$.

Exercício: demonstrar que

$$\|C_n(n\omega_0)\| = \|C_{-n}(-n\omega_0)\| = \frac{1}{2} \sqrt{a_n^2 + b_n^2} = \frac{1}{2} A_n, \quad (\text{B-0.9})$$

e que

$$\phi_n(n\omega_0) = \arctan\left(-\frac{b_n}{a_n}\right) = \phi_n, \quad (\text{B-0.10})$$

e

$$\phi_{-n}(n\omega_0) = \arctan\left(\frac{b_n}{a_n}\right) = -\phi_n. \quad (\text{B-0.11})$$

Exemplo 3: calcular o espectro $C_n(n\omega_0)$ da onda quadrada da figura B.1. Verificar o resultado utilizando as relações entre C_n , a_n e b_n .

Para a onda quadrada da figura B.1

$$C_n(n\omega_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} s(t) e^{-jn\omega_0 t} dt.$$

Integrando a exponencial nos mesmos intervalos que no exercício anterior obtemos

$$C_n(n\omega_0) = -\frac{1}{jn\omega_0 T_0} [2e^{-jn\omega_0 T_0/4} - 2e^{jn\omega_0 T_0/4} + e^{jn\omega_0 T_0/2} - e^{-jn\omega_0 T_0/2}],$$

utilizando a conhecida forma $\sin x = (1/2j)(e^{jx} - e^{-jx})$ obtemos a partir da relação anterior

$$C_n(n\omega_0) = \frac{2}{jn\omega_0 T_0} [2j \sin(n\omega_0 T_0/4)] - \frac{1}{jn\omega_0 T_0} [2j \sin(n\omega_0 T_0/2)];$$

simplificando por j e sabendo que o segundo termo é nulo temos que

$$C_n(n\omega_0) = \frac{\sin(n\omega_0 T_0/4)}{n\omega_0 T_0/4} = \frac{\sin n\pi/2}{n\pi/2}.$$

A figura B.3 mostra uma representação gráfica de $C_n(n\omega_0)$ para $-10 < n < 10$.

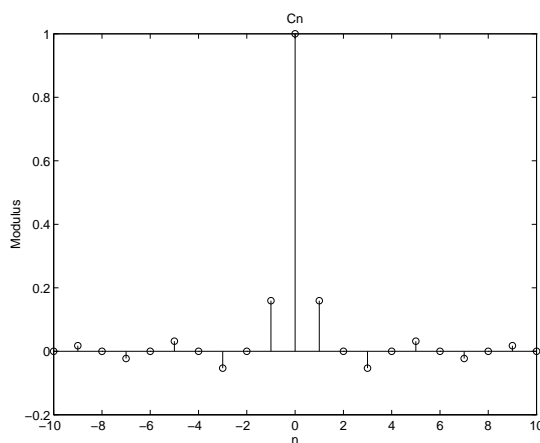


Figura B.3: espectro de amplitude discreto C_n .

Utilizando os valores de a_n e b_n do exercício anterior e sabendo que $|C_n| = 1/2\sqrt{a_n^2 + b_n^2}$ e que $\phi(n\omega_0) = \phi_n = 0$ obtem-se facilmente o resultado acima.

Exemplo 4: calcular o espectro $C_n(n\omega_0)$ do sinal da figura B.2. Aplicando a forma geral

$$C_n(n\omega_0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} \cos(\omega_0 t) e^{-jn\omega_0 t} dt.$$

Utilizando a forma exponencial do cosseno,

$$C_n(n\omega_0) = \frac{1}{2T_0} \int_{-T_0/2}^{T_0/2} e^{j\omega_0 t(1-n)} + e^{-j\omega_0 t(1+n)} dt,$$

integrando no devido intervalo e voltando a substituir a forma exponencial da função seno

$$C_n(n\omega_0) = \frac{1}{\omega_0 T_0} \left[\frac{1}{1-n} \sin\left[\omega_0(1-n)\frac{T_0}{2}\right] + \frac{1}{1+n} \sin\left[\omega_0(1+n)\frac{T_0}{2}\right] \right];$$

pondo sob a forma habitual de $\sin x/x$,

$$C_n(n\omega_0) = \frac{1}{2} \frac{\sin[\omega_0(1-n)\frac{T_0}{2}]}{\omega_0(1-n)\frac{T_0}{2}} + \frac{1}{2} \frac{\sin[\omega_0(1+n)\frac{T_0}{2}]}{\omega_0(1+n)\frac{T_0}{2}} = \frac{1}{2} \frac{\sin(1-n)\pi}{(1-n)\pi} + \frac{1}{2} \frac{\sin(n+1)\pi}{(n+1)\pi},$$

esta função é sempre nula para qualquer valor de n salvo para $n = -1$ e $n = 1$ onde toma o valor $1/2$. Podemos então desenhar o seu espectro de amplitude

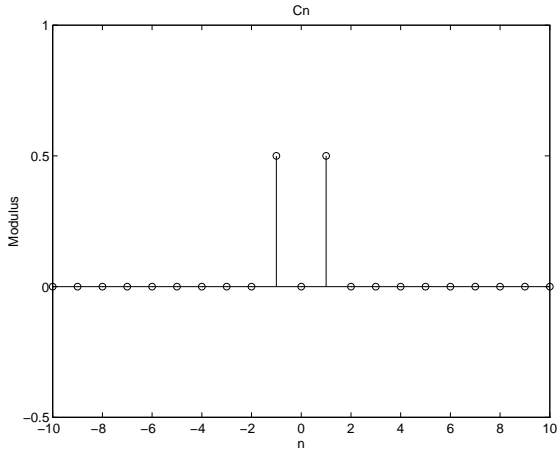


Figura B.4: espectro de amplitude discreto C_n de uma senoide.

C Noção de variável aleatória

Em muitos problemas físicos é necessário associar um valor numérico ao resultado de uma experiência aleatória. Por exemplo, o sinal da figura C.1 poderia representar a voltagem medida aos bornos de uma resistência geradora de ruído. Poderia também representar a temperatura, em graus centígrados, medida num dado local em função do tempo. Em ambos os casos ao resultado de uma experiência imprevisível foi associado um valor numérico. Este conceito, extremamente importante na modelização de sinais físicos não determinísticos, será progressivamente introduzido neste capítulo.

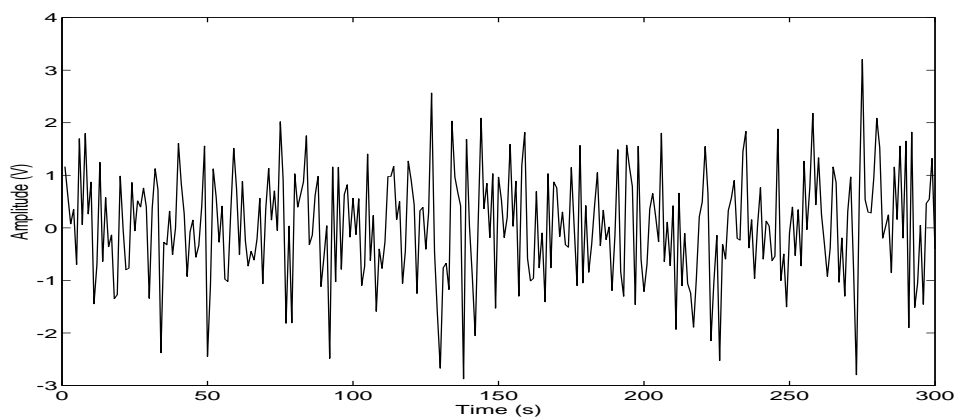


Figura C.1: sinal aleatório.

C.1 Probabilidades elementares

Começamos por considerar uma determinada experiência cujo resultado designaremos por ω . A realização dessa experiência pode resultar numa série, ou conjunto, de acontecimentos que constitui o conjunto de acontecimentos possíveis Ω . Para o experimentador, de entre os resultados possíveis pode ser assinalável o acontecimento de $\omega_0 \in \Omega$ para o qual gostaria de saber qual a “probabilidade” para que aconteça. Intuitivamente podemos dizer que a probabilidade do acontecimento ω_0 , que notaremos $Pr(\omega_0)$, pode ser considerada como o limite, quando o número de experiências tende para infinito, do quociente entre o número de vezes que ω_0 aconteceu e o número total de experiências realizadas. Assim,

$$Pr(\omega_0) = \lim_{N \rightarrow \infty} \frac{\text{numero de } \omega = \omega_0}{N} \quad (\text{C-1.1})$$

Portanto temos que $0 \leq Pr(\omega_0) \leq 1$. A probabilidade definida em C-1.1 é chamada probabilidade elementar do acontecimento ω_0 . Recordam-se algumas regras básicas de manipulação de probabilidades elementares.

A cada acontecimento, A , podemos associar o acontecimento complementar ou seja a não realização de A que notaremos A^* . Intuitivamente podemos dizer que $Pr(A) + Pr(A^*) = 1$. Considerando agora um outro acontecimento, B , cuja probabilidade é $Pr(B)$, podemos definir acontecimentos compostos, ou seja, um novo acontecimento composto pela realização de A e B , que nós notaremos $A \cap B$ ou AB . Para podermos deduzir qual o valor da probabilidade deste novo acontecimento AB temos que recordar o conceito de independência: dois acontecimentos A e B são independentes se a realização de um deles não altera em nada a probabilidade de realização do outro. Assim, se A e B são independentes $Pr(AB) = Pr(A)Pr(B)$. Outra forma de composição de acontecimentos é o de obter o acontecimento A ou o acontecimento B . Isto nota-se $A \cup B$ ou $A + B$. A probabilidade $Pr(A + B)$ é igual à probabilidade de A , $Pr(A)$, mais a probabilidade de B , $Pr(B)$, se A e B nunca acontecem simultaneamente, ou seja se $Pr(AB) = 0$, então $Pr(A + B) = Pr(A) + Pr(B)$. Neste caso diz-se que os acontecimentos A e B são mutuamente exclusivos.

Por último, devemos relembrar o conceito de probabilidade condicional que se pode introduzir considerando a probabilidade que um dado acontecimento A aconteça sabendo que B já aconteceu e que se escreve $Pr(A|B)$ (lê-se ‘probabilidade de A se B ’). Por definição

$$Pr(A|B) = \frac{Pr(AB)}{Pr(B)} \quad (\text{C-1.2})$$

Combinando $Pr(A|B)$ e $Pr(B|A)$ obtemos a regra de Bayes que se escreve

$$Pr(B|A) = \frac{Pr(B)Pr(A|B)}{Pr(A)} \quad (\text{C-1.3})$$

C.2 Conceito de variável aleatória e de função de distribuição

Teóricamente, uma variável aleatória (VA) é o resultado de uma aplicação do espaço de probabilidade Ω no eixo dos números reais $R[-\infty, \infty]$. Em termos práticos, trata-se apenas de uma função determinística $X(\omega)$, que à realização de um certo acontecimento aleatório ω associa um valor numérico $X(\omega)$. Esse valor numérico adquire assim um carácter aleatório porque depende de uma tiragem aleatória. A caracterização duma VA, $X(\omega)$, faz-se através da função de distribuição $F_X(x)$ definida por

$$F_X(x) = Pr[\omega|X(\omega) \leq x]. \quad (\text{C-2.1})$$

Para simplificar a notação é costume designar a VA $X(\omega)$, simplesmente por X e a sua função de distribuição por $F(x)$. Uma noção importante é a de igualdade entre VA's. Duas VA's X e Y são consideradas iguais com probabilidade 1 se

$$F(x) = F(y) \quad (\text{C-2.2})$$

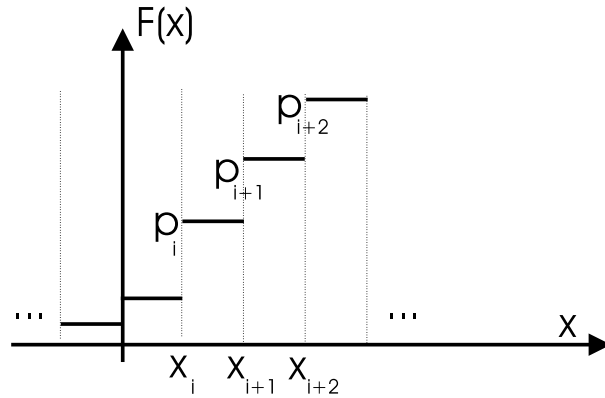


Figura C.2: função de distribuição de uma VA discreta.

ou seja, se possuem a mesma função de distribuição. Importa salientar que tal não significa que $X = Y$ em todos os pontos, mas apenas que o conjunto de pontos $I = \{\omega | X(\omega) = Y(\omega)\}$ é tal que $Pr(I) = 1$. A função de distribuição tem como propriedades evidentes

$$F(-\infty) = 0 \quad \text{e} \quad F(+\infty) = 1, \quad (\text{C-2.3})$$

e também que

$$Pr(a < X \leq b) = F(b) - F(a). \quad (\text{C-2.4})$$

Esta última relação indica que $F(b) - F(a) \geq 0$ o que significa também que a função de distribuição é uma função monótona não decrescente.

Até agora temos falado de VA's contínuas mas podemos, de igual forma, definir VA's discretas. Se o conjunto de acontecimentos possíveis Ω tem um número finito de elementos, e se considerarmos uma função $X(\omega)$ mensurável, então esta só poderá tomar um número finito de valores designados por $X_i = X(\omega_i)$. Neste caso X é uma VA discreta. A função de distribuição de uma VA discreta pode ser definida por

$$F(x) = \sum_{i=1}^k Pr(X_i = x_i), \quad x_k < x < x_{k+1} \quad (\text{C-2.5})$$

Como resultado, a função de distribuição de uma VA discreta é uma função crescente em escada, tomando valores constantes entre cada valor x_i, x_{i+1} como representado na figura 8.2.

C.3 Função densidade de probabilidade

Admitindo que a função de distribuição da VA contínua X é derivável, para qualquer x , podemos definir a função densidade de probabilidade $p(x)$ por

$$p(x) = \frac{dF(x)}{dx}, \quad (\text{C-3.1})$$

que admite como propriedades derivadas de (8.6) e (8.7),

$$F(b) - F(a) = \int_a^b p(x)dx \quad (\text{C-3.2})$$

e

$$\int_{-\infty}^{\infty} p(x)dx = 1, \quad (\text{C-3.3})$$

por outro lado usando (8.9)

$$p(x) = Pr(x < X \leq x + dx). \quad (\text{C-3.4})$$

Para as VA's discretas torna-se difícil definir a derivada de uma função do tipo da representada na figura 8.2 pois esta apresenta discontinuidades. Esta dificuldade pode ser contornada se considerarmos que a função de distribuição $F(x)$ de uma VA discreta como a representada na figura é o resultado de uma soma de funções degrau unidade atrasadas de x_i e multiplicadas por p_i , assim

$$F(x) = \sum_{i=1}^k p_i u(x - x_i), \quad (\text{C-3.5})$$

a partir da qual, e utilizando a “função” Delta de (x) , como a derivada da função degrau unidade, $u(x)$, obtemos

$$p(x) = \frac{dF(x)}{dx} = \sum_{i=1}^k p_i \delta(x - x_i). \quad (\text{C-3.6})$$

C.4 Esperança matemática

Acábamos de ver que uma VA X se encontra completamente caracterizada pela sua função de distribuição $F(x)$. Isto é um “luxo” que raramente acontece na prática, onde nos limitamos ao conhecimento do seu valor médio e/ou desvio padrão. O valor médio, ou esperança matemática de X é definida por

$$E[X] = \int xp(x)dx, \quad (\text{C-4.7})$$

que dá uma noção de média de X , no sentido de probabilidade, quando o número de experiências aumenta. Para uma VA discreta

$$E[X] = \sum x_i p_i. \quad (\text{C-4.8})$$

Da mesma maneira, a esperança matemática da função determinística da VA X , $g(x)$ é definida por

$$E[g(x)] = \int g(x)p(x)dx. \quad (\text{C-4.9})$$

Uma das mais importante formas que poderá tomar a função $g(x)$ será provavelmente x^k , o que permite introduzir a definição de momento de ordem k como sendo

$$m_k = E[X^k] = \int x^k p(x)dx. \quad (\text{C-4.10})$$

Para $k = 1$ obtém-se o momento de ordem um, ou média, definida em (8.15). Em muitos casos teremos $m_1 = 0$ porém, quando $m_1 \neq 0$ a VA $Y = X - m_1$ terá média nula, o que leva ao conceito de momentos centrados de ordem k cuja definição é

$$\mu_k = E[(X - m_1)^k]. \quad (\text{C-4.11})$$

O momento mais conhecido é μ_2 , a que se chama variância, e que é igual ao quadrado do desvio padrão σ^2

$$\mu_2 = \sigma^2 = E[(X - m_1)^2] = m_2 - m_1^2. \quad (\text{C-4.12})$$

Recordemos que o desvio padrão dá uma noção da variação da VA em torno ao seu valor médio.

C.5 Função característica

A função característica de uma VA X é definida pela esperança matemática da função $g(x) = \exp(juX)$, assim

$$\phi_X(u) = E[\exp(juX)] = \int \exp(jux)p(x)dx, \quad (\text{C-5.13})$$

e no caso em que X é uma VA discreta

$$\phi_X(u) = \sum_i \exp(jux_i)p_i. \quad (\text{C-5.14})$$

Uma das maiores utilidades da função característica é a forma relativamente simples que esta função toma para certas leis de distribuição que veremos mais adiante. Além disso, a partir da função característica é extremamente simples deduzir os momentos

de qualquer ordem. Uma relação muito útil é a que se obtém através da substituição de $\exp(juX)$ pelo seu desenvolvimento em série,

$$\begin{aligned} e^{jux} &= 1 + juX + \frac{(juX)^2}{2!} + \dots \\ &= \sum_{k=0}^{\infty} \frac{j^k u^k x^k}{k!} \end{aligned} \quad (\text{C-5.15})$$

que permite escrever

$$\phi_X(u) = \sum_{k=0}^{\infty} \frac{j^k u^k}{k!} E[x^k] \quad (\text{C-5.16})$$

onde o momento de ordem k é $m_k = E[x^k]$ e então

$$\phi_X(u) = \sum_{k=0}^{\infty} \frac{j^k u^k}{k!} m_k \quad (\text{C-5.17})$$

ou ainda derivando ambos os termos desta última equação em ordem a u , e tomando o resultado no ponto $u = 0$ obtemos

$$m_k = (-j)^k \frac{d^k \phi(u)}{du^k} \Big|_{u=0}. \quad (\text{C-5.18})$$

Diz-se que a função característica é a função geradora dos momentos.

C.6 Exemplos

C.6.1 Distribuição uniforme

A densidade de probabilidade de uma VA X que obedece a uma distribuição uniforme escreve-se

$$p(x) = \begin{cases} \frac{1}{b-a} & a \leq x \leq b \\ 0 & x \notin [a, b] \end{cases} \quad (\text{C-6.1})$$

a respectiva função de distribuição $F(x)$ é

$$F(x) = \begin{cases} \frac{x-a}{b-a} & a \leq x \leq b \\ 0 & x \leq a \\ 1 & b \leq x \end{cases} \quad (\text{C-6.2})$$

Exercício: demonstrar que a esperança matemática e a variância de uma VA uniforme se escrevem respectivamente

$$E[X] = \frac{a+b}{2} V[X] = \sigma^2 = \frac{(b-a)^2}{12} \quad (\text{C-6.3})$$

C.6.2 Distribuição exponencial unilateral

A densidade de probabilidade de uma VA X distribuída segundo uma lei exponencial escreve-se

$$p(x) = u(x)a \exp(-ax), \quad a > 0 \quad (\text{C-6.4})$$

onde $u(x)$ é a função degrau unidade definida anteriormente. A esperança matemática e variância são

$$E[X] = \frac{1}{a} V[X] = \sigma^2 = \frac{1}{a^2} \quad (\text{C-6.5})$$

Exercício: demonstrar a esperança matemática e variância acima. Calcular a função característica $\phi(u)$.

C.6.3 Distribuição exponencial bilateral

A densidade de probabilidade de uma VA X distribuída segundo uma lei exponencial bilateral escreve-se

$$p(x) = \frac{a}{2} \exp(-a|x|), \quad a > 0 \quad (\text{C-6.6})$$

que é evidentemente uma função par e portanto todos os momentos de ordem ímpar são nulos. Assim

$$E[X] = 0 \quad (\text{C-6.7})$$

$$V[X] = \sigma^2 = \frac{2}{a^2} \quad (\text{C-6.8})$$

Exercício: demonstrar a esperança matemática e variância.

C.6.4 Distribuição normal (ou gaussiana)

A densidade de probabilidade de uma VA normal X , algumas vezes notada abreviadamente por $N(m, \sigma^2)$, é dada por

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}, \quad (\text{C-6.9})$$

a sua média é evidentemente igual a m e variância igual a σ^2 .

Exercício: demonstrar que a função característica de uma variável aleatória gaussiana se escreve

$$\phi(u) = e^{jum - \frac{\sigma^2 u^2}{2}} \quad (\text{C-6.10})$$

C.6.5 Distribuição do χ^2 com d graus de liberdade

A densidade de probabilidade de uma VA X segundo a lei do χ^2 com d inteiro > 1 graus de liberdade escreve-se

$$p(x) = \begin{cases} 0 & x \leq 0 \\ \frac{e^{-\frac{x}{2}} x^{\frac{d-2}{2}}}{2^{d/2} \Gamma(\frac{d}{2})} & x \geq 0 \end{cases} \quad (\text{C-6.11})$$

onde

$$\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx \quad \text{para } x > 0 \quad (\text{C-6.12})$$

e cuja média é $m = d$ e o desvio padrão é $\sigma = \sqrt{2d}$. A importância desta lei de probabilidade vem do facto de que se tivermos n VA's X_1, X_2, \dots, X_n independentes e de mesma lei gaussiana $G(m, \sigma)$ então a VA

$$Y = \sum_{i=1}^n \left(\frac{X_i - m}{\sigma} \right)^2, \quad (\text{C-6.13})$$

segue uma lei de probabilidade $\chi^2(n)$. A expressão C-6.13 surge frequentemente quando se trabalha com sinais aleatórios (discretos) de densidade de probabilidade gaussiana e se pretende calcular a sua energia total num dado segmento de tempo (ou frequência). Nesse caso i representa o tempo (ou a frequência), n representa o tempo total e a soma dos quadrados C-6.13 não é mais senão a energia total do sinal no intervalo de tempo considerado.

D Folhas de Exercícios

D.1 Revisões sobre sistemas e sinais

Exercício 1: Um sistema é representado pela sua resposta impulsiva $h(t)$. A excitação $x(t)$ do sistema é dada por

$$x(t) = \begin{cases} (\frac{E}{T})t, & 0 < t < T/2; \\ E - (\frac{E}{T})t, & T/2 < t < T; \\ 0, & t > T. \end{cases}$$

- a) representar o sinal de excitação $x(t)$
- b) determinar por convolução, a resposta $y(t)$ à excitação $x(t)$.

Exercício 2:

- a) demonstre que se $F_1(s)$ e $F_2(s)$ forem respectivamente as Transformadas de Laplace dos sinais $f_1(t)$ e $f_2(t)$ então

$$\text{TL}[f_1(t) * f_2(t)] = F_1(s)F_2(s)$$

- b) demonstre que qualquer função $x(t)$ pode ser representada por

$$x(t) = \int_{-\infty}^{+\infty} x(\tau)\delta(t - \tau)dt$$

onde $\delta(t)$ é o Dirac.

Exercício 3:

Calcule as Transformadas de Fourier dos seguintes sinais

- a) $s_1(t) = \text{rect}(t)$
- b) $s_2(t) = \text{sinc}(t)$
- c) $s_3(t) = \exp[-(1/2)(t/\sigma)^2]$
- d) $s_4(t) = \delta(t)$

e) $s_5(t) = 1$

f) $s_6(t) = u(t)$

g) $s_7(t) = A \sin(\omega_0 t)$

Exercício 4:

Considere o sinal complexo

$$x(t) = A \exp(j\omega_0 t)$$

a) calcule a sua Transformada de Fourier

b) calcule a sua função de autocorrelação

c) calcule a sua densidade espectral de potência

D.2 Sinais determinísticos I

Exercício 1:

O sinal amostrado

$$\hat{x}(t) = \tau x(t) \sum_{n=-\infty}^{\infty} \delta(t - n\tau), \quad \text{com} \quad \tau < \pi/\omega_c$$

é reconstituído pelo seguinte filtro interpolador passa-baixo ideal

$$H_r(\omega) = H_0 e^{-jt_0\omega} p_{\omega_c}(\omega)$$

onde $p_{\omega_c}(\omega)$ é uma função porta de largura $2\omega_c$. Tendo em conta que o sinal $x(t)$ é de banda limitada, *i.e.*, $X(\omega) = 0$ para $|\omega| > \omega_c$, demonstre que a resposta do filtro interpolador é

$$y(t) = H_0 x(t - t_0)$$

Exercício 2:

Uma linha de atraso seguida de um integrador como indicado na figura D.1 é um exemplo de um dispositivo que “retem” o valor de uma amostra em memória e que pode servir para reconstituir o sinal contínuo a partir de um sinal amostrado.

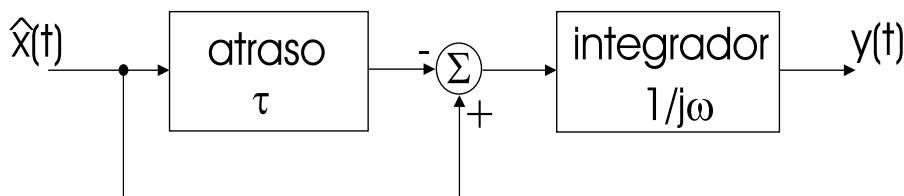


Figura D.1: circuito sample and hold.

a) demonstre que a função do sistema é

$$H(j\omega) = \frac{Y(j\omega)}{\hat{X}(j\omega)} = \tau \frac{\sin(\omega\tau/2)}{\omega\tau/2} e^{-\frac{j\omega\tau}{2}}$$

b) demonstre gráficamente o funcionamento deste circuito para um sinal de entrada amostrado com um intervalo de amostragem τ igual ao atraso da linha.

c) determine o espectro $Y(f)$ do sinal de saída em função do espectro inicial $X(f)$.
Note que

$$\hat{x}(t) = x(t) \sum_n \delta(t - n\tau)$$

Exercício 3:

Dado um sinal contínuo e periódico de período T_0 tal que

$$x(t) = \sum_{n=-\infty}^{\infty} C_n e^{jn\omega_0 t}$$

- a) calcular a sua TF
- b) aplicar ao cálculo da TF de um "trem de Diracs" $\Theta(f)$

$$\gamma(t) = \sum_{k=-\infty}^{\infty} \delta(t - kT)$$

Exercício 4:

Demonstre que se um sistema de resposta impulsiva complexa tem uma entrada real ele pode ser representado por dois sistemas de resposta impulsiva real. Faça um esboço desse sistema. Demonstre que o mesmo se aplica a um sistema real com uma entrada complexa.

Exercício 5:

- a) demonstre que um sistema complexo pode ser representado por quatro sistemas reais.
- b) demonstre que se os quatro sistemas reais que representam um sistema complexo forem lineares o sistema complexo também é linear.

Exercício 6:

Dada a representação de um sinal amostrado $\hat{x}(t)$ função de uma soma de impulsos modulados em amplitude, demonstre que

$$\hat{X}(f) = \text{TF}[\hat{x}(t)] = \text{TF}[x(k)]$$

D.3 Sinais determinísticos II

Exercício 1:

Um sistema "phase splitter" é tal que a sua resposta em frequência se escreve

$$\Phi(\omega) = \begin{cases} 1, & \omega \geq 0 \\ 0, & \omega < 0 \end{cases}$$

demonstre que a resposta do sistema "phase splitter" a um sinal real $x(t)$ é

$$y(t) = \frac{1}{2}\{x(t) + j\hat{x}(t)\}$$

onde $\hat{x}(t) = H[x(t)]$ é a transformada de Hilbert de $x(t)$.

Exercício 2:

Considere o esquema de blocos da figura D.2, onde $y(t)$ é um sinal passa banda com o seu espectro centrado em ω_c e $\phi(t)$ é a resposta impulsiva de um "phase splitter" indicado no exercício 1.

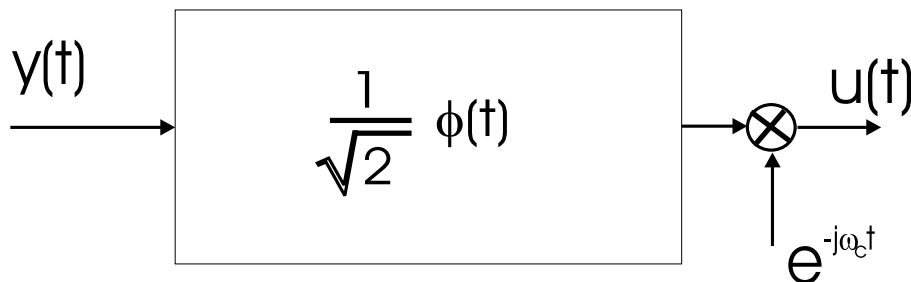


Figura D.2: phase splitter

a) demonstre que o sinal de saída $u(t)$ é um sinal passa baixo que se escreve

$$u(t) = \frac{1}{\sqrt{2}}\{y(t) + j\hat{y}(t)\}e^{-j\omega_c t}$$

b) demonstre que $u(t)$ e $y(t)$ têm a mesma energia devido ao coeficiente $\sqrt{2}$.

Exercício 3:

Atendendo a que $H[x(t)] = \hat{x}(t)$ demonstre que

$$H[\hat{x}(t)] = -x(t)$$

Exercício 4:

Considerando o sinal

$$x(t) = \cos(\omega_0 t)$$

- a) calcule a transformada de Hilbert $\hat{x}(t)$
- b) calcule $H[\hat{x}(t)]$. Verifique o resultado do exercício anterior.

Exercício 5:

- a) calcule a transformada de Hilbert $\hat{x}(t)$ de

$$x(t) = \frac{1}{1+t^2}$$

- b) faça um esboço de $x(t)$ e $\hat{x}(t)$

Exercício 6:

Considere o espectro de um sistema realizável $h(t)$

$$H(\omega) = R(\omega) + jX(\omega)$$

com

$$R(\omega) = \pi\delta(\omega)$$

Determine a parte imaginária $X(\omega)$

D.4 Sinais aleatórios I

Exercício 1:

Considere uma VA X , gaussiana de função densidade de probabilidade (FDP)

$$p(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-m)^2}{2\sigma^2}\right\}$$

com os parâmetros $m = 0$ e $\sigma^2 = 1$. Demonstre que

- o integral de $p(x)$ é igual a 1
- a sua variância é também igual a 1.
- a sua função característica $\Phi_X(u)$ é

$$\log \Phi_X(u) = jmu + \frac{\sigma^2 u^2}{2}$$

Exercício 2:

Ao resultado da experiência de jogar uma moeda ao ar associamos uma VA discreta X . Esta VA discreta só toma um número finito de valores, neste caso igual a dois: “cara” ou “coroa”. À acontecimento de obter coroa associamos a probabilidade p , $Pr(\omega = \text{' coroa' } | X = 0) = p$, assim $Pr(\omega = \text{' cara' } | X = 1) = 1 - p = q$.

- qual a esperança matemática $E[X]$? E o momento de ordem k , $E[X^k]$?
- calcular a variância $V[X]$.
- demonstrar que a função característica da VA X , $\phi_X(u)$ se escreve

$$\phi_X(u) = 1 + q[\exp(ju) - 1]$$

Exercício 3:

Uma VA discreta X segundo a distribuição de Poisson toma os valores inteiros $0, 1, 2, \dots$ com as probabilidades,

$$p_k = Pr(X = k) = \frac{m^k}{k!} \exp(-m) \quad (1)$$

- demonstrar que o momento de ordem 1, $m_1 = m$.

b) demonstrar que o momento de ordem 2, $\sigma^2 = m$.

c) calcular as probabilidades

$$P_+ = Pr(X = \text{numero par})$$

e

$$P_- = Pr(X = \text{numero impar})$$

sabendo que, obviamente, $P_+ + P_- = 1$.

d) calcular a função característica de X , $\phi_X(u)$ com a distribuição de Poisson (1).

Exercício 4:

Demonstrar que para uma VA Gaussiana X de média μ e variância σ^2 temos

$$Pr[X > x] = Q\left(\frac{x - \mu}{\sigma}\right)$$

Exercício 5:

Considere uma combinação linear arbitrária de N VA's Gaussianas X_i , independentes, de média zero e variância σ^2 ,

$$Z = a_1 X_1 + \dots + a_N X_N$$

Utilizando a função característica demonstre que Z é também Gaussiana de média nula e de variância

$$\sigma_Z^2 = (a_1^2 + \dots + a_N^2)\sigma^2$$

D.5 Sinais aleatórios II

Exercício 1:

Demonstre que um processo estocástico branco e estacionário X_k , filtrado por um filtro de resposta impulsiva h_k já não é branco mas continua a ser estacionário.

Exercício 2:

A densidade de probabilidade de Cauchy é

$$p(x) = \frac{a/\pi}{x^2 + a^2} \quad -\infty < x < \infty \quad (1)$$

- a) determine a média e a variância de X
- b) determine a função característica de X

Exercício 3:

Uma VA Y é definida por

$$Y = \frac{1}{n} \sum_{i=1}^n X_i$$

onde $X_i; i = 1, \dots, n$ é um conjunto de n VA's estatisticamente independentes e indênticamente distribuídas segundo a distribuição de Cauchy (1).

- a) determine a função característica de Y
- b) determine a densidade de probabilidade de Y
- c) considere a densidade de probabilidade de Y quando $n \rightarrow \infty$. O teorema do limite central verifica-se? Justifique a sua resposta.

D.6 Sinais aleatórios III

Exercício 1:

Considere um processo estocástico com a seguinte forma

$$X(t) = \sum_{n=-\infty}^{\infty} a_n g(t - nT)$$

onde $\{a_n\}$ é uma sequência discreta de variáveis aleatórias de média $m_n = E[a_n]$ e função de autocorrelação

$$r_{aa}(k) = \frac{1}{2} E[a_n a_{n+k}^*],$$

que representa a mensagem a transmitir e $g(t)$ é um sinal determinístico que representa a função de pulso.

Calcule

- a) a média do processo $X(t)$
- b) a função de autocorrelação de $X(t)$, $r_{xx}(t + \tau, t)$
- c) demonstre que $X(t)$ é um processo cicloestacionário
- d) para $m_a = 0$, $r_{aa}(k) = \frac{\sigma_a^2}{2} \delta(k)$ e

$$g(t) = \begin{cases} \cos(\omega_0 t), & -T/2 \leq t \leq T/2, \omega_0 = \pi/T \\ 0, & \text{outro } t \end{cases}$$

determine $r_{xx}(t + \tau, t)$.

e) nas mesmas condições que em d) determine a função de autocorrelação média do processo $X(t)$

$$\bar{r}_{xx}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} r_{xx}(t + \tau, t) dt$$

f) a partir de $\bar{r}_{xx}(\tau)$ calculada em e) determine a densidade espectral média de potência, $\bar{P}_{xx}(f) = \text{TF}[\bar{r}_{xx}(\tau)]$.

Exercício 2:

Considere um processo estocástico de média nula e estacionário $X(t)$ com a densidade espectral de potência

$$P_{xx}(f) = \begin{cases} 1, & |f| \leq W \\ 0, & |f| > W \end{cases}$$

O processo $X(t)$ é amostrado a uma taxa $1/T$ para obter um processo discreto $X(n) = X(nT)$ determine:

- a expressão da função de autocorrelação de $X(n)$
- o valor mínimo de T que resulta numa sequência branca e constante na frequência
- repita b) para uma densidade espectral de potência de $X(t)$

$$P_{xx}(f) = \begin{cases} 1 - |f|/W, & |f| \leq W \\ 0, & |f| > W \end{cases}$$

Exercício 3:

A função de autocorrelação de um processo estocástico de ruído branco $X(t)$ é

$$r_{xx}(\tau) = \frac{1}{2}N_0\delta(\tau)$$

Supondo que $x(t)$ é o sinal de entrada de um sistema tendo como resposta em frequência

$$|H(f)| = \begin{cases} 1, & -B/2 \leq |f - f_c| \leq B/2 \\ 0, & \text{outro valor de } f \end{cases}$$

Determine a potência total de ruído à saída do filtro.

D.7 Modulação e desmodulação I

Exercício 1:

Considere o seguinte sinal modulador

$$m(t) = \begin{cases} \text{sinc}(100t), & |t| \leq t_0 \\ 0, & \text{outro } t \end{cases}$$

com $t_0 = 0.1$. Este sinal modula uma portadora de frequência $f_c = 250$ Hz em AM-SC.

- represente graficamente o sinal modulador $m(t)$
- calcule o sinal modulado $u(t)$
- determine e represente graficamente os espectros de $m(t)$ e de $u(t)$.
- determinar a potência do sinal modulador e do sinal modulado

Exercício 2:

Utilizando o mesmo sinal modulador do exercício anterior determinar:

- a transformada de Hilbert do sinal modulador $m(t)$ no domínio da frequência
- o sinal $u(t)$ modulado em AM-LSSB e o seu espectro.
- a potência do sinal modulador e do sinal modulado.

D.8 Modulação e desmodulação II

Exercício 1:

Seja $v_i(t)$ e $v_q(t)$ dois sinais passa baixo numa banda $W < f_c$, com energias E_i e E_q respectivamente. Utilize a generalização do teorema de Parseval (equação de Rayleigh)

$$\int_{-\infty}^{\infty} v(t)w^*(t)dt = \int_{-\infty}^{\infty} V(f)W^*(f)df$$

para demonstrar que

a) $\int_{-\infty}^{\infty} v_{bp}(t)dt = 0$ onde o sinal passa banda

$$v_{bp}(t) = v_i(t) \cos(\omega_c t) - V_q(t) \sin(\omega_c t)$$

e que

b) a energia do sinal passa banda é igual a $(E_i + E_q)/2$

Exercício 2:

Considere o sinal modulador $x(t) = \cos(2\pi f_c t)u(t)$. Represente gráficamente o sinal modulado no tempo $x_c(t)$ e na frequência $X_c(f)$ para

- a) uma modulação AM com $a = 1$
- b) uma modulação AM com $a > 1$
- c) uma modulação DSB

Exercício 3:

Provar que para o sinal modulado

$$x_c(t) = A_c[1 + ax(t)] \cos(\omega_c t + \phi)$$

onde $x(t)$ é um sinal aleatório ergódico de média nula e ϕ é uma variável aleatória de fase, independente de $x(t)$ e uniformemente distribuída em $[0, 2\pi]$ a energia média escreve-se

$$E[x_c^2(t)] = \frac{1}{2}A_c^2(1 + a^2S_x)$$

onde $S_x = E[x^2(t)]$.

D.9 Modulação PAM I

Exercício 1:

Considere um processo estocástico $Z(t)$, definido por

$$Z(t) = S(t)e^{j\omega_c t}$$

onde $S(t)$ é um processo estocásticos estacionário. Demonstre que se $E[S(t)] = 0$ então $Z(t)$ é também estacionário e que as funções de correlação são ligadas por

$$R_Z(\tau) = R_S(\tau)e^{j\omega_c \tau}$$

Determine igualmente o espectro $P_Z(\omega)$ de $Z(t)$.

Exercício 2:

Demonstre que $Z(t)$ e $Z^*(t)$ (do exercício anterior) são conjuntamente estacionários se e só se

$$R_{SS^*}(\tau) = 0$$

Demonstre ainda que esta mesma condição implica também que

$$R_{ZZ^*}(\tau) = 0$$

Exercício 3:

Considerando o sinal estocástico complexo $S(t) = R(t) + jI(t)$, demonstre que $R_{SS^*}(\tau) = 0$ implica

$$R_R(\tau) = R_I(\tau)$$

e que

$$R_{RI}(\tau) = -R_{IR}(\tau) = -R_{RI}(\tau)$$

Exercício 4:

O equivalente em banda base de um sinal PAM passa banda pode-se escrever

$$S(t) = \sum_{k=-\infty}^{\infty} A_k h(t - kT + \Theta)$$

onde $h(t)$ pode ser complexo e Θ é um termo de fase aleatório. Considerando que A_k é uma sequência aleatória estacionária e independente de Θ , demonstre que

a) uma condição suficiente para que $R_{SS^*}(\tau) = 0$ é que

$$E[A_k A_m] = 0 \quad (1)$$

b) que para que (1) seja satisfeita é suficiente que as partes real e imaginária de A_k tenham a mesma função de autocorrelação e que seja decorreladas uma da outra.

Exercício 5:

Seja

$$X(t) = \sqrt{2}\text{Re}\{Z(t)\} = \frac{1}{\sqrt{2}}[Z(t) + z^*(t)]$$

a) demonstre que $X(t)$ é estacionário se e só se $S(t)$ tiver média nula, for também estacionário e tal que

$$R_{SS^*}(\tau) = 0$$

b) demonstre também que sob as condições enumeradas em a) a função de autocorrelação de $X(t)$ se escreve

$$R_X(\tau) = \text{Re}\{R_Z(\tau)\} = \text{Re}\{e^{j\omega_c\tau} R_S(\tau)\}$$

c) calcule a densidade espectral de $X(t)$.

D.10 Modulação PAM II

Exercício 1:

Considere um sistema PAM em banda base que utiliza impulsos raised cosine. Fazendo a hipótese de que a sequência a transmitir é branca e normalizada de forma que o seu espectro se escreve

$$S_A(\omega) = 1$$

demonstre que a potência transmitida é independente de T para qualquer valor do factor de roll-off α .

Exercício 2:

Considere um canal limitado a $|\omega/2\pi| \leq 1500$ Hz. Qual é o valor máximo da symbol rate que pode ser atingida nesse canal para um excesso de banda de 50% o filtro de recepção é do tipo passa-baixo e que não existe ISI.

Exercício 3:

Considere o seguinte sinal PAM em banda base

$$u(t) = \sum_n [a_n g(t - 2nT) - j b_n g(t - 2nT - T)]$$

onde $\{a_n\}$ e $\{b_n\}$ são duas sequências aleatórias estatisticamente independentes e a função de pulso $g(t)$ é

$$g(t) = \begin{cases} \sin(\pi t/2T), & 0 < t < 2T \\ 0, & \text{outro } t \end{cases}$$

No sinal $u(t)$ as sequências $\{a_n\}$ e $\{b_n\}$ são transmitidas à velocidade $1/2T$ bits/s enquanto $u(t)$ é transmitido a $1/T$ bits/s.

a) demonstre que o envelope $|u(t)|$ é constante independentemente de $\{a_n\}$ e de $\{b_n\}$.

b) determine a densidade espectral de $u(t)$

Exercício 4:

Considere um sinal 4-PSK definido pelo seu equivalente banda base

$$u(t) = \sum_n I_n g(t - nT)$$

onde I_n toma um dos valores entre quatro possíveis $\{1/\sqrt{2}(\pm 1, \pm j)\}$ com igual probabilidade. A sequência resultante é branca.

a) calcule e represente a densidade espectral de $u(t)$ quando

$$g(t) = \begin{cases} A, & 0 \leq t \leq T \\ 0, & \text{outro } t \end{cases}$$

b) repita a) com

$$g(t) = \begin{cases} A \sin(\pi t/T), & 0 \leq t \leq T \\ 0, & \text{outro } t \end{cases}$$

c) compare os espectros obtidos em a) e b) em termos de largura de banda a -3 dB e largura de banda no primeiro zero.

Referências

- [1] Edward A. Lee and David G. Messerschmidt *Digital Communication*, Sec. Edition, Kluwer Academic Publishers, 1994.
- [2] John G. Proakis, *Digital Communications*, Third Ed., McGraw-Hill, New York, 1995.
- [3] John G. Proakis and Masoud Salehi, *Comtemporary Communications Systems using MATLAB*, Brooks Cole, Pacific Grove, 2000.
- [4] M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions*. U.S. Government Printing Office, 1968.

E Método de avaliação

- Mini-testes de avaliação: MT1, MT2 e MT3.
- Práticas (P): média da avaliação contínua nas aulas práticas.

Nota de controle contínuo CC

$$CC = \frac{MT1 + MT2 + MT3 + P}{4}$$

Só serão admitidos a exame os alunos que tenham uma assiduidade mínima nas aulas práticas de 80% e uma média P igual ou superior a 10 (dez).