

Introdução à sincronização em modulações digitais

© Sílvio A. Abrantes

Departamento de Engenharia Electrotécnica e de Computadores
Faculdade de Engenharia, Universidade do Porto
Porto, Portugal
sam@fe.up.pt

Março de 2007
(revisto em Abril de 2009)

“Sincronização: acto ou efeito de sincronizar ou de sincronizar-se;
Sincronizar: fazer coincidir no tempo”
Dos dicionários

1. Introdução

A sincronização em qualquer das modalidades referidas mais adiante é algo que habitualmente se tem como adquirido: transmite-se um sinal numa dada modulação – seja QAM, FSK, GMSK ou outra qualquer – e considera-se que, por arte “mágica”, fácil ou difícil, há sincronização perfeita: a onda de referência do correlacionador tem a fase e a frequência da portadora recebida, o intervalo de integração do integrador começa e acaba nos instantes correctos, enfim, considera-se que todas essas operações de “cronómetro” foram executadas com precisão! Mas... de facto, como se chega a essa boa situação? Como conhecer a fase da portadora? Como cronometrar os símbolos?

A desmodulação de um sinal modulado digitalmente é realizada de modo *coerente* ou *não-coerente*. Na desmodulação coerente é usado um filtro adaptado ou o seu equivalente, o correlacionador da Fig. 1, onde T é a duração dos símbolos da sequência de informação enviada. O correlacionador, como é sabido, realiza duas operações: 1) multiplica o sinal de entrada por uma onda sinusoidal de referência que deve ter a mesma fase e frequência da portadora modulada recebida; 2) efectua uma integração que deve começar quando cada símbolo de informação começa e terminar quando ele acaba. A saída do correlacionador é a sequência dos valores atingidos no fim de cada período de integração, de T em T segundos.

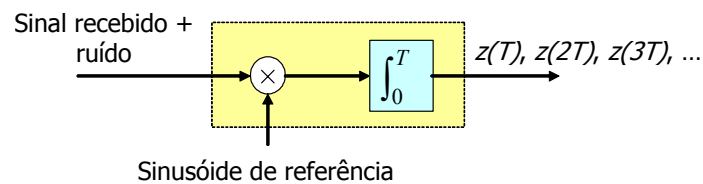


Fig. 1 O correlacionador da desmodulação coerente.

Se a sinusóide de referência possuir a mesma fase e frequência da portadora ambas as ondas se cruzam por zero ao mesmo tempo, como mostra a Fig. 2.

À tarefa de gerar a onda local sincronizada dá-se o nome de *sincronização da fase* ou *recuperação da fase* da portadora. À tarefa de determinar os instantes de início e fim de cada símbolo dá-se o nome de *sincronização de símbolos*, *recuperação da temporização de símbolos* ou *recuperação de relógio*. São estes os temas a abordar neste texto.

A necessidade de uma correcta recuperação de relógio é ilustrada em seguida recorrendo a filtros adaptados. Imaginemos que à entrada de um filtro adaptado é aplicado um impulso rectangular como o da Fig. 3a, à esquerda. A saída do filtro é o impulso triangular da mesma Fig. 3a, à direita, o qual deve ser amostrado no instante em que a amplitude é máxima, ou seja, no instante $t = T$ segundos. A amostragem em instante diferente, por exemplo em $t = T - \delta$ ou em $t = T + \delta$, origina um valor inferior ao valor de pico. A Fig. 3c mostra a saída do mesmo filtro se à entrada se apresentar a sequência de impulsos rectangulares da

Fig. 3b, entretanto recebida com um atraso τ . Um relógio mal sincronizado (adiantado, na figura) origina um conjunto de amostras de valores inferiores aos ideais, o que aumenta a probabilidade de erro.

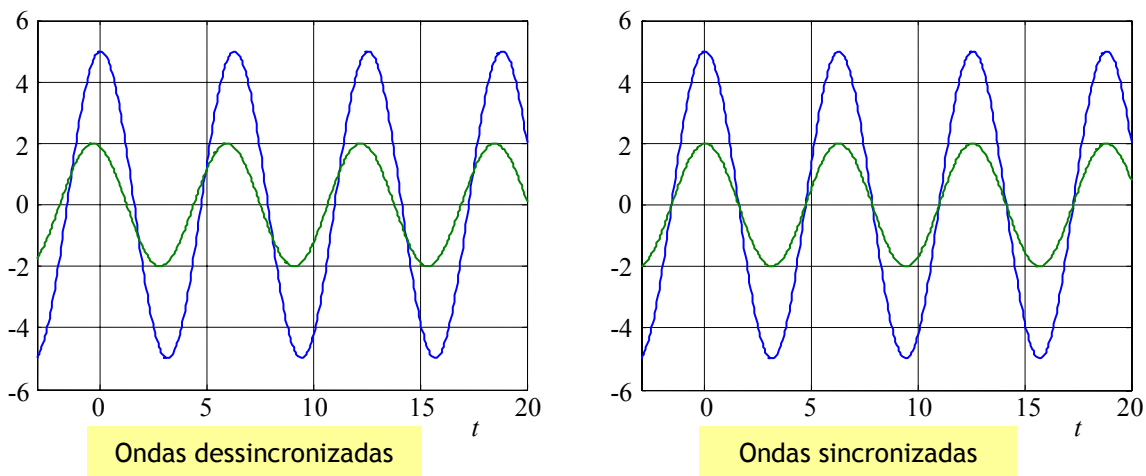


Fig. 2 Ondas sincronizadas e dessincronizadas em fase.

Com um correlacionador a situação é semelhante. O intervalo de integração adequado vai de 0 a T segundos (como na Fig. 4 e na Fig. 5a); se, pelo contrário, a integração for realizada durante um período diferente, como entre $-\delta$ e $T-\delta$ ou entre δ e $T+\delta$, os valores de saída são indesejavelmente menores. Por exemplo, na Fig. 5b a sequência de saída do correlacionador, representada pelas bolas vermelhas, pode tomar os valores ± 1 e $\pm(1-2\delta/T)$ em vez de apenas ± 1 . Isto, é claro, tem consequências na probabilidade de erro.

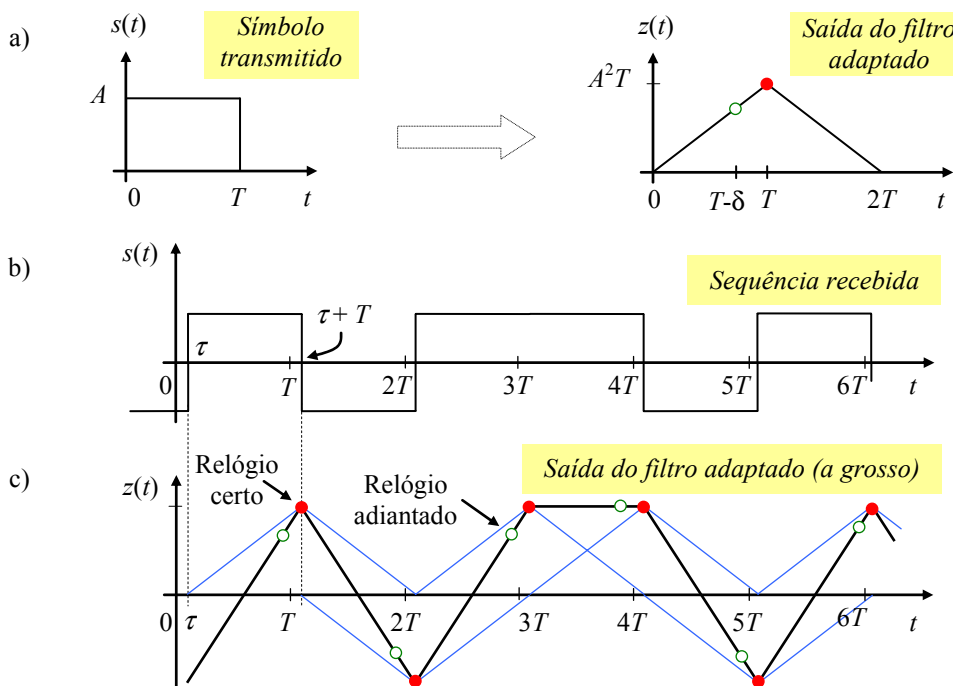


Fig. 3 Resposta do filtro adaptado a impulsos rectangulares.

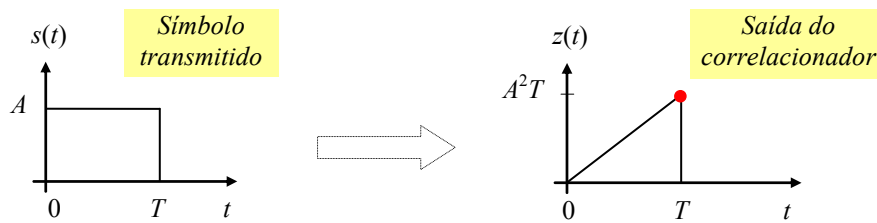


Fig. 4 Resposta do correlacionador a um impulso rectangular. A integração decorre de 0 a T segundos.

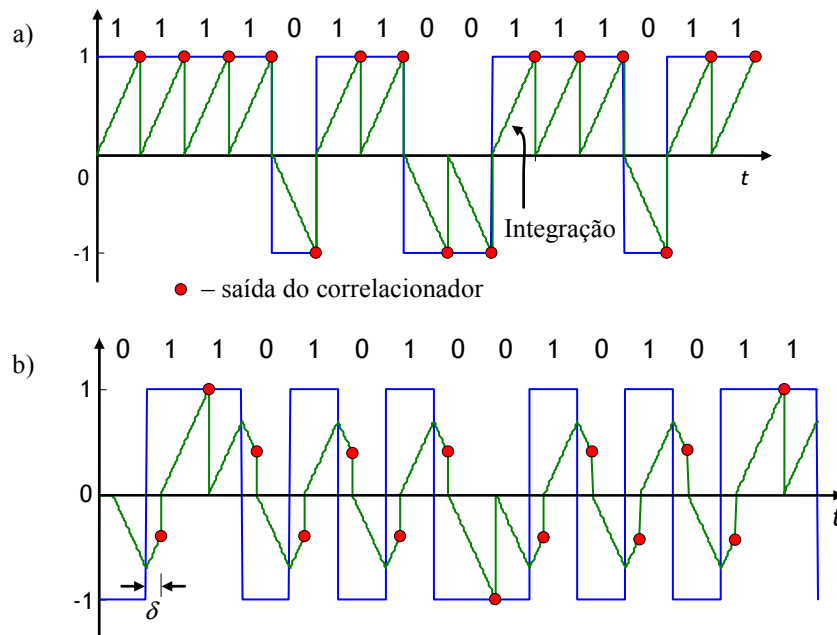


Fig. 5 Resposta do correlacionador a sequências rectangulares antipodais. a) temporização correcta; b) temporização incorrecta (a integração começa e acaba tarde demais).

Em conclusão: é sempre importante estimar com precisão o início de cada símbolo.

A recuperação de fase da portadora e a recuperação da temporização de símbolo podem ser realizadas numa operação conjunta ou em operações separadas. Neste texto seguiremos apenas a segunda opção.

Além das sincronizações de fase e de símbolo pode ser também necessário proceder à *sincronização de trama* – para, numa sequência dividida em tramas, se poder saber quando cada uma começa e acaba – ou realizar a *sincronização de blocos de símbolos* – para, numa situação de correcção de erros com códigos de blocos, se poder saber quando começam e acabam as palavras de código.

Vale a pena referir que as tarefas de sincronização podem ser auxiliadas se, juntamente com o sinal, for transmitida informação de sincronismo adequada, informação que pode revestir várias formas. Assim, para auxiliar a recuperação da fase e frequência da portadora pode ser transmitida também uma *portadora-piloto*, não modulada – desperdiçando, é claro, potência; para estimar o início e o fim de um símbolo, de uma trama ou de uma palavra de código podem ser transmitidos *bits de sincronismo* ou enviadas sequências pré-definidas de símbolos para “treinar” o receptor no início da transmissão – à custa, de novo, da eficiência da comunicação. As técnicas de sincronização “ajudadas pelos dados” acabadas de referir – portadoras-piloto, sequências de treino, etc. – são designadas na literatura anglófona por técnicas “Data-Aided” (DA) e têm em comum o facto de serem ajudas que o emissor envia ao receptor. Há técnicas de sincronização em que, pelo contrário, é o próprio receptor que se auto-ajuda, quando ele tem em conta as suas próprias decisões. Este processo, denominado de recuperação “Decision-Directed” (DD), pode ser interpretado como “Data-Aided”. Se não houver qualquer ajuda a sincronização diz-se “Non-Data-Aided” (NDA).

Só vamos considerar modulações digitais lineares (MPSK e QAM). Antes, porém, de vermos como nessas modulações a fase da portadora e a temporização dos símbolos podem ser estimados apreciemos as consequências dos erros de sincronismo nas constelações de sinal e na probabilidade de erro.

2. Efeito dos erros de sincronismo na probabilidade de erro

Suponhamos que uma sequência binária de informação com débito $R_b = 1/T$ bits/s é transmitida em PSK binário. A forma de onda genérica é $s(t) = b(t)p(t)\cos 2\pi f_c t$, $0 \leq t \leq T$, em que f_c é a frequência da portadora, $b(t) = \pm 1$, $0 \leq t \leq T$, é um bit da mensagem e $p(t)$ representa o impulso de formatação (impulso rectangular, impulso de cosseno elevado, etc.). Se for rectangular então $p(t) = 1$, $0 \leq t \leq T$, como nas figuras anteriores¹.

Ao ser transmitido, o sinal $s(t)$ sofre alterações ao longo do canal de comunicação. Por um lado, chega ao receptor com um atraso, τ , relativamente ao momento em que foi enviado; por outro, devido ao canal chega ao receptor atenuado de A e com um esfasamento, θ , desconhecido; finalmente, pode também sofrer algum indesejável desvio de frequência, f_d , no oscilador, em relação à frequência nominal da portadora. Tendo em conta todas estas modificações e também o ruído no canal o receptor recebe a forma de onda

$$\begin{aligned} r(t) &= Ab(t-\tau)\cos[2\pi(f_c + f_d)(t-\tau) + \theta] + n(t) = \\ &= Ab(t-\tau)\cos[2\pi(f_c + f_d)t + \underbrace{\theta - 2\pi(f_c + f_d)\tau}_{\phi}] + n(t) \end{aligned}$$

onde se fez $\phi = \theta - 2\pi(f_c + f_d)\tau$ e se considera

$$A = \sqrt{2E_b/T} \text{ – amplitude do sinal recebido } (E_b: \text{ energia do bit})$$

θ – esfasamento introduzido pelo canal

τ – atraso de propagação

f_d – desvio (“offset”) de frequência, no oscilador, em relação à frequência nominal da portadora

$n(t)$ – ruído aditivo gaussiano branco de densidade espectral de potência $N_0/2$.

Vamos admitir que a frequência f_c e o “offset” f_d são conhecidos. É tarefa do receptor coerente estimar com a precisão possível os valores do atraso τ e da fase ϕ . As suas estimativas são, respectivamente, $\hat{\tau}$ e $\hat{\phi}$. Quando estas estão incorrectas o desempenho do receptor piora, como vamos ver seguidamente.

2.1 Erros de fase em MPSK e QAM

Vamos supor que a recuperação da fase da portadora apresenta um erro de fase ϕ e que a temporização de símbolos é feita adequadamente. Começamos por PSK binário e depois prosseguiremos com MPSK e QAM. Em todos os casos a detecção é coerente e o ruído é gaussiano branco $\mathcal{N}(0, \sigma^2)$, com $\sigma^2 = N_0/2$.

BPSK

A Fig. 6a apresenta o diagrama de blocos simplificado de um detector coerente de BPSK tendo à entrada o sinal $\pm\sqrt{2E_b/T}\cos 2\pi f_c t$, $0 \leq t \leq T$. A onda local de referência é $\sqrt{2/T}\cos(2\pi f_c t + \phi)$. Efectuando as operações de multiplicação e integração indicadas conclui-se que com o erro de fase ϕ a constelação do sinal à entrada do decisor é composta de dois pontos situados em $\pm\sqrt{E_b}\cos\phi$ (Fig. 6b). Sem erro os pontos estariam situados em $\pm\sqrt{E_b}$, mais longe um do outro, portanto, o que mostra que o erro de fase aproxima os pontos da constelação e com isso aumenta a probabilidade de bit errado de $P_b = Q(\sqrt{2E_b/N_0})$ para

¹ Neste texto vamos considerar apenas impulsos de formatação rectangulares.

$$P_b = Q\left(\sqrt{2E_b/N_0} \cos \phi\right)$$

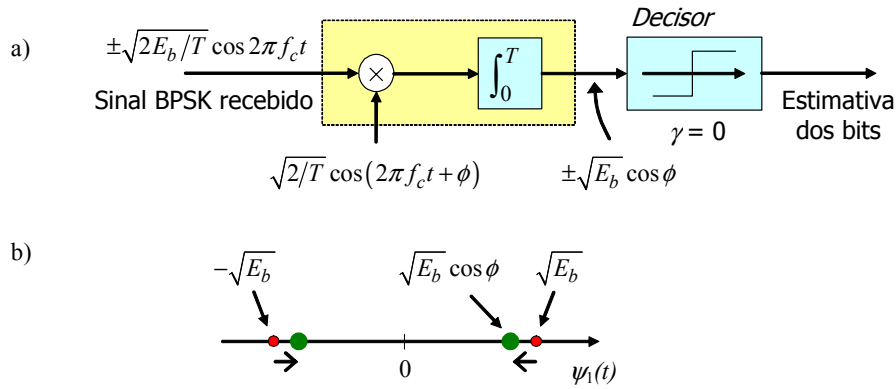


Fig. 6 Desmodulação coerente em BPSK com erro de fase.
a) diagrama de blocos do detector; b) constelação.

MPSK

O efeito do erro de fase em MPSK ainda é pior. Sendo E e T a energia e a duração de cada símbolo, respectivamente, o sinal é dado em $0 \leq t \leq T$ por

$$\begin{aligned} s_i(t) &= \sqrt{2E/T} \cos(2\pi f_c t - \phi_i) = \\ &= \left(\sqrt{2E/T} \cos \phi_i\right) \cos 2\pi f_c t + \left(\sqrt{2E/T} \sin \phi_i\right) \sin 2\pi f_c t \end{aligned}$$

com $\phi_i = \frac{2\pi}{M}i$, $i=1, \dots, M$, a que corresponde o ponto genérico $\mathbf{s}_i = \left[\sqrt{E} \cos \phi_i \quad \sqrt{E} \sin \phi_i\right]^T$ na representação em espaço de sinal bidimensional. Em vez de ser multiplicado por $\sqrt{2/T} \cos(2\pi f_c t)$ e $\sqrt{2/T} \sin(2\pi f_c t)$ o sinal é multiplicado pelas referências $\sqrt{2/T} \cos(2\pi f_c t + \phi)$ e $\sqrt{2/T} \sin(2\pi f_c t + \phi)$. De acordo com a Fig. 7, à entrada do decisor vamos encontrar os dois valores seguintes, correspondentes, respectivamente, aos ramos em fase e em quadratura:

$$\begin{aligned} z_{i1} &= \sqrt{E} \cos \phi_i \cos \phi - \sqrt{E} \sin \phi_i \sin \phi = \\ &= \sqrt{E} \cos(\phi_i + \phi) \\ z_{i2} &= \sqrt{E} \cos \phi_i \sin \phi + \sqrt{E} \sin \phi_i \sin \phi = \\ &= \sqrt{E} \sin(\phi_i + \phi) \end{aligned}$$

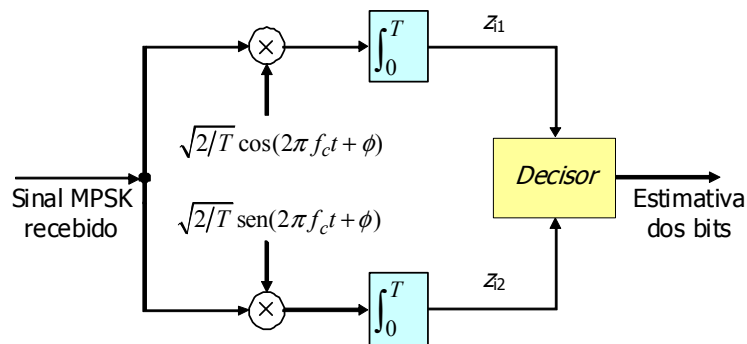


Fig. 7 Desmodulação coerente de MPSK com erro de fase.

Os valores z_{i1} e z_{i2} são as coordenadas do sinal recebido no espaço de sinal, $\mathbf{z}_i = [z_{i1} \ z_{i2}]^T$, os quais, sem erro de fase, valeriam $\sqrt{E} \cos \phi_i$ e $\sqrt{E} \sin \phi_i$. Ou seja, o erro de fase provoca uma rotação dos pontos da constelação de um ângulo ϕ , rotação que pode ser expressa em notação vectorial se desenvolvermos as expressões de z_{i1} e z_{i2} :

$$\mathbf{z}_i = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \begin{bmatrix} \sqrt{E} \cos \phi_i \\ \sqrt{E} \sin \phi_i \end{bmatrix} = \begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix} \mathbf{s}_i.$$

O que isto mostra é que a matriz $\begin{bmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{bmatrix}$ faz rodar o ponto \mathbf{s}_i de um ângulo ϕ no sentido contrário aos ponteiros do relógio. A uma rotação assim obtida dá-se o nome de *rotação de Givens*.

Chegaríamos às mesmas conclusões se usássemos notação complexa: sendo cada ponto da constelação representado pelo complexo $\mathbf{s}_i = \sqrt{E} \cos \phi_i + j\sqrt{E} \sin \phi_i$ então o ponto recebido é

$$\mathbf{z}_i = z_{i1} + jz_{i2} = \mathbf{s}_i e^{j\phi}.$$

Mais uma vez vemos que o ponto \mathbf{s}_i rodou de um ângulo ϕ . A Fig. 8 mostra-o para QPSK, onde $\phi_i = \pi/4, 3\pi/4, 5\pi/4, 7\pi/4$. O ponto do primeiro quadrante, em particular, tem coordenadas

$$z_{11} = \sqrt{E/2}(\cos \phi - \sin \phi)$$

$$z_{12} = \sqrt{E/2}(\cos \phi + \sin \phi)$$

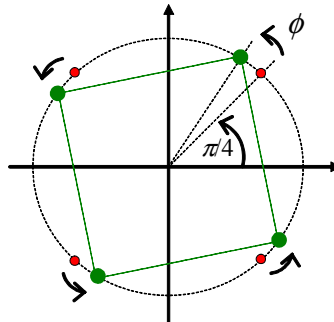


Fig. 8 Rotação da constelação QPSK devido ao erro de fase ϕ .

A rotação da constelação tem implicações na probabilidade de erro, que aumenta, visto que as regiões de decisão – os quadrantes – não mudam². Analisemos a questão em QPSK, primeiro, e depois em MPSK com $M \geq 4$. A probabilidade de símbolo errado, P_e , é igual a um menos a probabilidade de o símbolo estar certo, P_c :

$$P_e = 1 - P_c.$$

Mas $P_c = \sum_{i=1}^4 p_i P_{c_i}$, em que p_i é a probabilidade de ocorrência de cada símbolo e P_{c_i} é a probabilidade condicional de decisão correcta dado se ter transmitido o símbolo de ordem i , probabilidade que é igual ao produto das probabilidades de decisão correcta segundo cada eixo. Ora, dada a simetria da constelação todos os P_{c_i} são iguais pelo que $P_c = P_{c_i}$, $\forall i$. Tomando de novo como exemplo o ponto do primeiro quadrante, as

² A invariância à rotação na probabilidade de erro aqui não se aplica pois as regiões de decisão não rodam com a constelação.

probabilidades de erro segundo os eixos das abcissas e das ordenadas são $Q\left(\frac{z_{11}}{\sigma}\right)$ e $Q\left(\frac{z_{12}}{\sigma}\right)$, respectivamente, pelo que

$$P_{ci} = \left[1 - Q\left(\frac{z_{11}}{\sigma}\right)\right] \left[1 - Q\left(\frac{z_{12}}{\sigma}\right)\right] = 1 - Q\left(\frac{z_{11}}{\sigma}\right) - Q\left(\frac{z_{12}}{\sigma}\right) + Q\left(\frac{z_{11}}{\sigma}\right)Q\left(\frac{z_{12}}{\sigma}\right)$$

Desprezando o produto de funções Q por ser muito pequeno temos então, para a probabilidade de símbolo errado:

$$P_e \approx Q\left(\frac{z_{11}}{\sigma}\right) + Q\left(\frac{z_{12}}{\sigma}\right) = Q\left[\sqrt{E/N_0}(\cos\phi - \text{sen}\phi)\right] + Q\left[\sqrt{E/N_0}(\cos\phi + \text{sen}\phi)\right]$$

Mas em QPSK a energia de símbolo é o dobro da energia de bit, $E = 2E_b$, e a probabilidade de bit errado é metade da probabilidade de símbolo errado, $P_b = P_e/2$, com mapeamento de Gray. Logo, finalmente,

$$P_b \approx \frac{1}{2}Q\left[\sqrt{\frac{2E_b}{N_0}}(\cos\phi - \text{sen}\phi)\right] + \frac{1}{2}Q\left[\sqrt{\frac{2E_b}{N_0}}(\cos\phi + \text{sen}\phi)\right]$$

A representação gráfica de P_b é feita na Fig. 9 para vários valores do erro de fase.

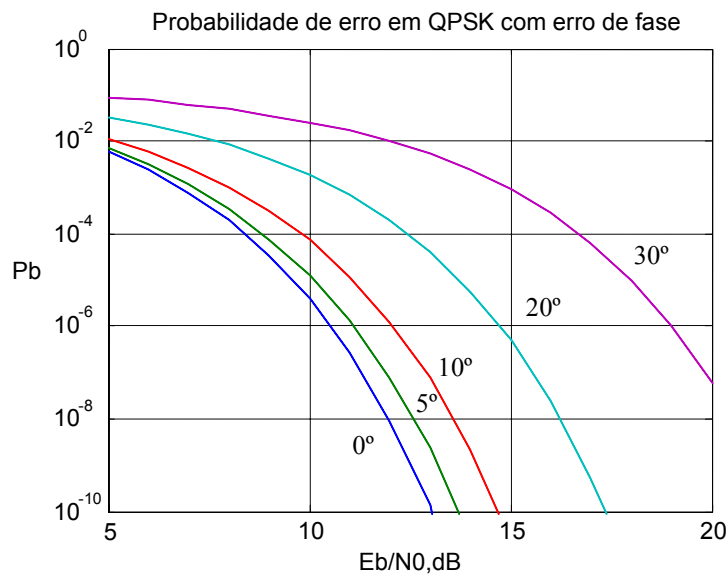


Fig. 9 Probabilidade de bit errado em QPSK com erro de fase na portadora local.

Vejamos agora como generalizar as conclusões anteriores para uma constelação MPSK com $M \geq 4$. Em primeiro lugar observemos a Fig. 10, onde se mostra a região de decisão de um dado ponto da constelação rodado de um ângulo ϕ . O ponto recebido está à distância A de uma fronteira de decisão e à distância B da outra.

Para valores elevados da relação E/N_0 a probabilidade de símbolo errado em MPSK é estimada pela aproximação dos vizinhos mais próximos $P_e \approx \sum_{\substack{k=1 \\ k \neq i}}^2 Q\left(\frac{d_{ik}}{\sqrt{2N_0}}\right)$ (tendo em conta que em MPSK cada ponto tem

2 vizinhos assim). O parâmetro d_{ik} representa, é claro, a distância do ponto de ordem i ao ponto de ordem k , ou o dobro da distância à fronteira de decisão. Adaptando a expressão à Fig.10 ficamos com

$$P_e \approx Q\left(\frac{2A}{\sqrt{2N_0}}\right) + Q\left(\frac{2B}{\sqrt{2N_0}}\right), \text{ em que}$$

$$A = \sqrt{E} \text{sen}(\pi/M - \phi)$$

$$B = \sqrt{E} \text{sen}(\pi/M + \phi)$$

e \sqrt{E} é a distância dos pontos da constelação à origem. Agora é só substituir A e B na expressão da probabilidade de símbolo errado:

$$P_e \approx Q\left(\sqrt{\frac{2E}{N_0}} \text{sen}(\pi/M - \phi)\right) + Q\left(\sqrt{\frac{2E}{N_0}} \text{sen}(\pi/M + \phi)\right).$$

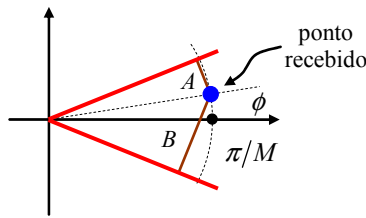


Fig. 10 Região de decisão de um ponto MPSK rodado de um ângulo ϕ .
O ponto está à distância \sqrt{E} da origem.

Sabemos que $E = E_b \log_2 M$ e que com mapeamento de Gray entre símbolos e bits se verifica $P_b \approx P_e / \log_2 M$. Assim, a probabilidade de bit errado vem dada por

$$P_b \approx \frac{1}{\log_2 M} Q\left(\sqrt{2 \log_2 M} \frac{E_b}{N_0} \text{sen}(\pi/M - \phi)\right) + \frac{1}{\log_2 M} Q\left(\sqrt{2 \log_2 M} \frac{E_b}{N_0} \text{sen}(\pi/M + \phi)\right).$$

A expressão de QPSK apresentada atrás é, evidentemente, um caso particular desta fórmula mais geral.

QAM

Em QAM o erro de fase também é mais inconveniente que em PSK binário. O sinal QAM é dado por

$$s_i(t) = a_i \sqrt{2E_0/T} \cos 2\pi f_c t + b_i \sqrt{2E_0/T} \text{sen} 2\pi f_c t, \quad 0 \leq t \leq T$$

em que $a_i = \pm 1, \pm 3, \dots$, $b_i = \pm 1, \pm 3, \dots$ e $\mathbf{s}_i = [a_i \sqrt{E_0} \quad b_i \sqrt{E_0}]^T$ representa um ponto genérico da constelação. Se o sinal for multiplicado pelas referências locais nos ramos em fase e em quadratura $\sqrt{2/T} \cos(2\pi f_c t + \phi)$ e $\sqrt{2/T} \text{sen}(2\pi f_c t + \phi)$ as saídas dos dois integradores do receptor coerente valem

$$z_{i1} = a_i \sqrt{E_0} \cos \phi - b_i \sqrt{E_0} \text{sen} \phi \quad (\text{em fase})$$

$$z_{i2} = a_i \sqrt{E_0} \text{sen} \phi + b_i \sqrt{E_0} \cos \phi \quad (\text{em quadratura})$$

O que vemos? Vemos que, além da diminuição de amplitude por um factor $\cos \phi$, o sinal desejado – $a_i \sqrt{E_0}$ ou $b_i \sqrt{E_0}$ – também é afectado de “crosstalk” proporcional a $\text{sen} \phi$ proveniente do outro ramo do detector.

Como vimos atrás, os valores z_{i1} e z_{i2} são as coordenadas no espaço de sinal do vector de sinal recebido, $\mathbf{z}_i = [z_{i1} \ z_{i2}]^T$, quando se transmite o sinal $\mathbf{s}_i = [a_i\sqrt{E_0} \ b_i\sqrt{E_0}]^T$. Acontece que o vector \mathbf{z}_i pode ser obtido através da seguinte multiplicação:

$$\mathbf{z}_i = \begin{bmatrix} \cos \phi & -\text{sen } \phi \\ \text{sen } \phi & \cos \phi \end{bmatrix} \begin{bmatrix} a_i\sqrt{E_0} \\ b_i\sqrt{E_0} \end{bmatrix} = \begin{bmatrix} \cos \phi & -\text{sen } \phi \\ \text{sen } \phi & \cos \phi \end{bmatrix} \mathbf{s}_i$$

Ora isto representa, como sabemos, uma rotação de Givens: o vector \mathbf{z}_i é obtido rodando \mathbf{s}_i de um ângulo ϕ no sentido contrário ao dos ponteiros do relógio, como aconteceu na Fig. 11. Ou seja, tal como em MPSK, também em QAM um erro de fase ϕ faz rodar a constelação de um ângulo ϕ . Aliás, se usássemos a notação complexa $\mathbf{z}_i = z_{i1} + jz_{i2}$ e $\mathbf{s}_i = a_i\sqrt{E_0} + jb_i\sqrt{E_0}$ concluiríamos o mesmo pois $\mathbf{z}_i = \mathbf{s}_i e^{j\phi}$.

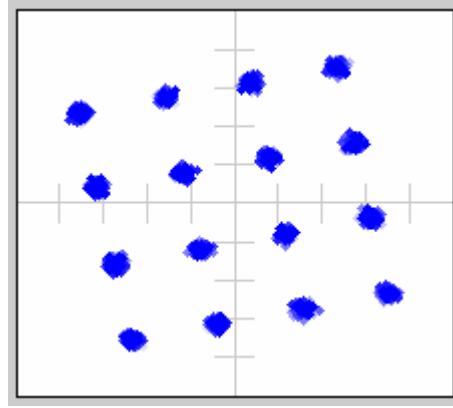


Fig. 11 Uma constelação 16-QAM com ruído rodada devido ao erro de fase.

A probabilidade de símbolo errado com erro de fase fixo ϕ pode ser calculada por um método semelhante ao usado em QPSK, identificando na constelação quadrada de M-QAM os $(\sqrt{M}-2)^2$ pontos interiores, os quatro pontos dos cantos e os $4\sqrt{M}-8$ pontos das bordas e calculando as probabilidades associadas. No entanto, não é preciso considerar toda a constelação – o que seria fastidioso, no mínimo – pois basta analisar apenas um dos quadrantes dado que a situação é igual em todos eles. Mesmo desprezando produtos cruzados de funções Q , como se fez em QPSK, a expressão final de P_e a que se chega não é simples [1]:

$$P_e \approx \frac{4}{M} \sum_{i=-\frac{\sqrt{M}}{2}+1}^{\frac{\sqrt{M}}{2}-1} \sum_{k=-\frac{\sqrt{M}}{2}+1}^{\frac{\sqrt{M}}{2}} Q \left\{ \sqrt{\frac{3}{M-1} \frac{E}{N_0}} [2i + (1-2i)\cos \phi + (2k-1)\text{sen } \phi] \right\}.$$

Válida para constelações quadradas com $M \geq 4$ e onde $E = \frac{2(M-1)E_0}{3}$ é a energia média dos símbolos, a expressão de P_e contém $\sqrt{M}(\sqrt{M}-1)$ parcelas com a função Q , um número bastante elevado (por exemplo, 12 parcelas em 16-QAM e 56 em 64-QAM!).

Usando mapeamento de Gray e sendo $E = E_b \log_2 M$ a probabilidade de bit errado vale

$$P_b \approx \frac{4/M}{\log_2 M} \sum_{i=-\frac{\sqrt{M}}{2}+1}^{\frac{\sqrt{M}}{2}-1} \sum_{k=-\frac{\sqrt{M}}{2}+1}^{\frac{\sqrt{M}}{2}} Q \left\{ \sqrt{\frac{3 \log_2 M}{M-1} \frac{E_b}{N_0}} [2i + (1-2i)\cos \phi + (2k-1)\text{sen } \phi] \right\}$$

No caso particular de 16-QAM a expressão escreve-se como

$$P_b \approx \frac{1}{16} \sum_{i=1}^1 \left[\sum_{k=-1}^2 Q \left\{ \sqrt{\frac{4 E_b}{5 N_0}} [2i + (1-2i) \cos \phi + (2k-1) \sin \phi] \right\} \right]$$

É claro que esta expressão engloba a inexistência de erro de fase. Na verdade, se $\phi = 0$ então

$$P_b \approx \frac{3}{4} Q \left(\sqrt{\frac{4 E_b}{5 N_0}} \right), \text{ que é precisamente a fórmula com sincronização perfeita.}$$

A Fig. 12 mostra a degradação em P_b provocada pelo erro de fase em 16-QAM.

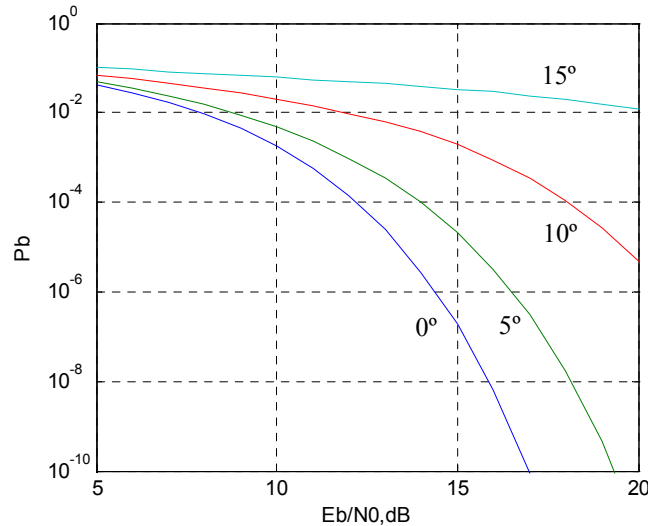


Fig. 12 Probabilidade de bit errado em 16-QAM com erro de fase na portadora local.

No exemplo seguinte vamos confirmar analiticamente que um mesmo erro de fase é mais grave em QPSK que em BPSK.

Exemplo 1:

Seja $E_b/N_0 = 9,6$ dB, para o qual $P_b \approx 10^{-5}$ com sincronização perfeita quer em BPSK quer em QPSK. Quais são as novas probabilidades de bit errado se o erro de fase for $\phi = 10^\circ$?

R.: Aplicando as expressões

$$\text{BPSK: } P_b = Q \left(\sqrt{\frac{2E_b}{N_0}} \cos \phi \right)$$

$$\text{QPSK: } P_b \approx \frac{1}{2} Q \left[\sqrt{\frac{2E_b}{N_0}} (\cos \phi - \sin \phi) \right] + \frac{1}{2} Q \left[\sqrt{\frac{2E_b}{N_0}} (\cos \phi + \sin \phi) \right]$$

e sabendo que 9,6 dB corresponde à razão 9,12 concluímos imediatamente que

$$\text{BPSK: } P_b = Q \left(\sqrt{2 \times 9,12} \cos \pi/18 \right) = 1,3 \cdot 10^{-5}$$

$$\begin{aligned} \text{QPSK: } P_b &\approx \frac{1}{2} Q \left[\sqrt{2 \times 9,12} (\cos \pi/18 - \sin \pi/18) \right] + \frac{1}{2} Q \left[\sqrt{2 \times 9,12} (\cos \pi/18 + \sin \pi/18) \right] = \\ &= 1,3 \cdot 10^{-4} \end{aligned}$$

Em QPSK o efeito de um mesmo erro de fase é, de facto, mais nefasto que em BPSK.

2.1.1 Resumo de expressões da probabilidade de erro quando há erro de fase ϕ

$$\text{BPSK: } P_b = Q\left(\sqrt{2E_b/N_0} \cos \phi\right)$$

$$\text{MPSK: } P_e \approx Q\left(\sqrt{2 \log_2 M \frac{E_b}{N_0}} \sin(\pi/M - \phi)\right) + Q\left(\sqrt{2 \log_2 M \frac{E_b}{N_0}} \sin(\pi/M + \phi)\right)$$

$$\text{QPSK: } P_e \approx Q\left[\sqrt{\frac{2E_b}{N_0}}(\cos \phi - \sin \phi)\right] + Q\left[\sqrt{\frac{2E_b}{N_0}}(\cos \phi + \sin \phi)\right].$$

$$\text{M-QAM: } P_e \approx \frac{4}{M} \sum_{i=-\frac{\sqrt{M}}{2}+1}^{\frac{\sqrt{M}}{2}-1} \left[\sum_{k=-\frac{\sqrt{M}}{2}+1}^{\frac{\sqrt{M}}{2}} Q\left\{ \sqrt{\frac{3 \log_2 M}{M-1} \frac{E_b}{N_0}} [2i + (1-2i) \cos \phi + (2k-1) \sin \phi] \right\} \right]$$

$$\text{16-QAM: } P_e \approx \frac{1}{4} \sum_{i=-1}^1 \left[\sum_{k=-1}^2 Q\left\{ \sqrt{\frac{4}{5} \frac{E_b}{N_0}} [2i + (1-2i) \cos \phi + (2k-1) \sin \phi] \right\} \right]$$

Com mapeamento de Gray e relação sinal-ruído elevada é sempre $P_b \approx P_e / \log_2 M$.

2.2 Erros de temporização em BPSK

Vamos supor agora que a recuperação da portadora é feita adequadamente e que a sincronização de símbolo apresenta um erro de δ segundos (ver Fig. 5b). Vimos que, na ausência de ruído, a sequência de saída do correlacionador tem quatro valores possíveis e equiprováveis, ± 1 e $\pm(1-2\delta/T)$ – ou melhor, $\pm(1-2|\delta|/T)$, para δ positivo ou negativo. Estabelecendo um limiar de decisão nulo não é difícil provar que a probabilidade de bit errado aumenta de $P_b = Q\left(\sqrt{2E_b/N_0}\right)$, quando não há erro de temporização, para

$$P_b = \frac{1}{2} Q\left(\sqrt{2E_b/N_0}\right) + \frac{1}{2} Q\left[\sqrt{2E_b/N_0} (1-2|\delta|/T)\right].$$

A Fig. 13 apresenta a probabilidade de bit errado em função do erro de temporização para três valores de E_b/N_0 : 8,4 dB, 9,6 dB e 10,5 dB (os valores com os quais se atingem as probabilidades 10^{-4} , 10^{-5} e 10^{-6} , respectivamente, quando $\delta=0$). Note-se a elevada degradação da probabilidade de erro mesmo para valores baixos de δ .

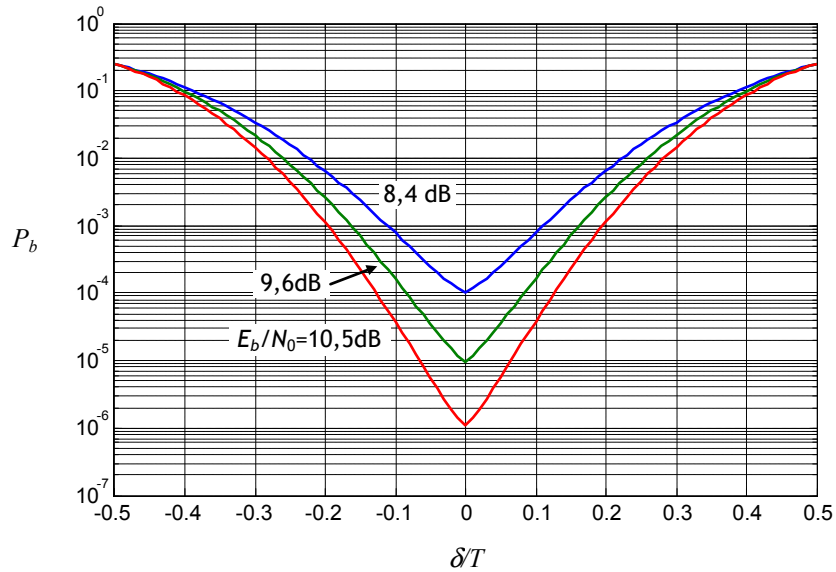


Fig. 13 Probabilidade de bit errado em BPSK em função do erro de temporização.

Exemplo 2:

Se no receptor coerente de BPSK dispusermos de $E_b/N_0 = 9,6$ dB a probabilidade de bit errado vale cerca de $P_b = 10^{-5}$ com sincronização perfeita, como vimos. Supondo que a recuperação da fase da portadora é exacta qual é a nova probabilidade de bit errado se existir um erro de temporização correspondente a 1% do tempo de bit?

R.: Substituindo valores na expressão de P_b obtemos imediatamente

$$P_b = \frac{1}{2} Q(\sqrt{2 \times 9,12}) + \frac{1}{2} Q[\sqrt{2 \times 9,12} (1 - 2 \times 0,01)] = 1,2 \cdot 10^{-5}$$

Um erro de 1% fez aumentar a probabilidade de erro 20%.

2.3 Esquemas de sincronização de fase e de símbolo

A sincronização “Non-Data-Aided” pode ser realizada com *malha fechada* ou *aberta*, consoante seja usada uma malha de realimentação ou não. A sincronização com *malha fechada* é a mais antiga das duas e, habitualmente, recorre a um PLL (“Phase-Locked Loop”). Por seu lado a sincronização “Decision-Directed” requer sempre uma malha fechada dado que é necessário proceder à realimentação das decisões.

A recuperação de fase com realimentação é normalmente realizada ou com um circuito denominado de “Costas loop” ou com um dispositivo não-linear que eleva o sinal a uma dada potência M . Quer um quer o outro removem a modulação e os seus desempenhos são considerados equivalentes. No que diz respeito à recuperação de relógio, sem realimentação usa-se muitas vezes um dispositivo quadrático, em tempo contínuo, e com realimentação um circuito chamado “Early-Late”, em tempo discreto. Uma alternativa a este circuito é o muito comum método de Gardner.

A recuperação de fase e a recuperação de temporização com malha aberta são tratadas, muitas vezes, como um problema de estimação de parâmetros; dois critérios são habitualmente considerados: o critério de estimação da *máxima verosimilhança* (ou ML, de “maximum likelihood”) e o critério de estimação da *máxima probabilidade a posteriori* (MAP, de “maximum a posteriori probability”).

Estes vários esquemas estão contemplados na Fig. 14.

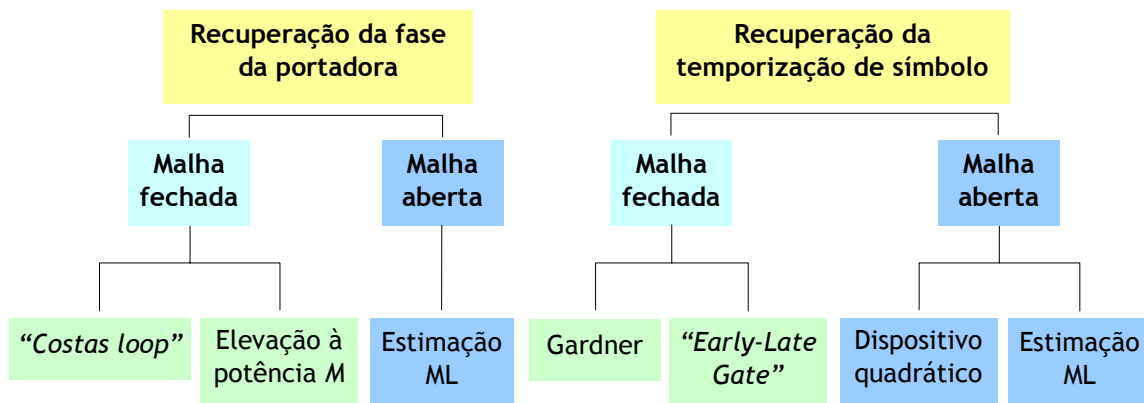


Fig. 14 Esquemas de recuperação da fase da portadora e da temporização de símbolo.

Os métodos de sincronização podem ainda dividir-se em métodos em tempo contínuo (de índole mais analógica) e em tempo discreto, ou amostrado. Vamos começar pelos primeiros.

3. Recuperação da fase da portadora em tempo contínuo

Vamos supor que não há “offset” de frequência ($f_d=0$) e que os símbolos estão perfeitamente sincronizados. Tomando, para simplificar, $\tau=0$ então $\phi=\theta$. Como estimar θ ? Começamos pela abordagem clássica, a sincronização com malha fechada, com realimentação.

3.1 Sincronização de fase com malha fechada

Nesta abordagem usa-se um de dois circuitos: ou o “Costas loop” ou um circuito não linear que eleva o sinal a uma potência M . Como em qualquer deles se utiliza um PLL vejamos como este funciona.

3.1.1 Funcionamento do PLL

Como mostra a Fig. 15, um PLL é constituído por três blocos: um *Detector de Fase*, um *Filtro de Malha* (“Loop Filter”) e um *VCO* (“Voltage-Controlled Oscillator”) [2]. O detector de fase é um *multiplicador* cujas entradas são formas de onda sinusoidais com a mesma frequência mas esfasadas de 90° – por outras palavras, se numa das entradas tivermos um cosseno, na outra teremos um seno, e vice-versa. A *frequência de repouso* (ou frequência não controlada) do VCO – aquela que é gerada quando a tensão de entrada é nula – deve ser, em funcionamento normal, igual ou muito próxima da frequência da onda de entrada do PLL. Finalmente, o filtro de malha é um filtro passa-baixo. Consoante as características deste assim o PLL é classificado como *de primeira ordem*, *de segunda ordem*, etc.

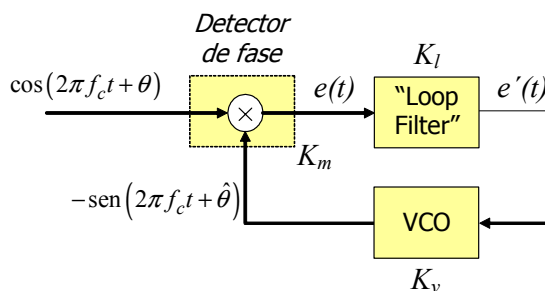


Fig. 15 Diagrama de blocos de um PLL.

Consideremos então que na entrada do PLL temos a sinusóide $\cos(2\pi f_c t + \theta)$. O objectivo do PLL é igualar a fase $\hat{\theta}$ da saída do VCO à fase do sinal de entrada, θ . Para o conseguir o circuito vai reduzindo a tensão do sinal de entrada do VCO até, eventualmente, a anular, o que é feito durante dois períodos de tempo

consecutivos: no primeiro, chamado *modo de aquisição*, $\hat{\theta}$ vai-se aproximando de θ ; depois, no segundo, no *modo de seguimento* (“tracking”), o PLL tenta manter o valor $\hat{\theta} \approx \theta$ adquirido no modo anterior³.

O detector de fase é realizado como um multiplicador de sinais. Estando a saída do VCO esfasada de 90° relativamente à entrada do detector de fase, $\cos(2\pi f_c t + \hat{\theta} + \pi/2) = -\sin(2\pi f_c t + \hat{\theta})$, a saída do detector vale⁴

$$\begin{aligned} e(t) &= -K_m \cos(2\pi f_c t + \theta) \sin(2\pi f_c t + \hat{\theta}) = \\ &= \frac{K_m}{2} [\sin(\theta - \hat{\theta}) - \sin(2\pi \cdot 2f_c t + \theta + \hat{\theta})] \end{aligned}$$

em que K_m é o ganho do multiplicador, em V^{-1} .

O filtro de malha é um filtro passa-baixo que atenua fortemente a componente em $2f_c$ de $e(t)$. Se se tratar de um filtro passa-baixo ideal de ganho constante, K_l , na sua banda de passagem o PLL diz-se de *primeira ordem*. Nesse caso,

$$e'(t) = \frac{K_l K_m}{2} \sin(\theta - \hat{\theta})$$

Na prática o ângulo $\theta - \hat{\theta}$ é muito pequeno (especialmente no modo de seguimento) pelo que podemos usar a aproximação $\sin x \approx x$, ou seja,

$$e'(t) \approx \frac{K_l K_m}{2} (\theta - \hat{\theta}).$$

A frequência angular instantânea, ω_i , da saída do VCO é igual à derivada do ângulo $2\pi f_c t + \hat{\theta}(t) = \omega_c t + \hat{\theta}(t)$:

$$\omega_i = \frac{d}{dt} [\omega_c t + \hat{\theta}(t)] = \omega_c + \frac{d\hat{\theta}(t)}{dt}.$$

Em tempo contínuo a fase de saída do VCO está relacionada com a tensão de entrada através da equação diferencial

$$\frac{d\hat{\theta}(t)}{dt} = K_v e'(t),$$

em que K_v é o ganho do VCO, em $\text{rad}/(V \cdot s)$ (ver a Fig. 16). Em tempo discreto a fase do VCO rege-se por

$$\hat{\theta}_{k+1} = \hat{\theta}_k + K_v e'_k,$$

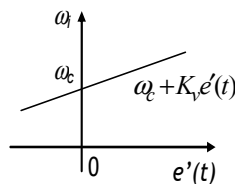


Fig. 16 Frequência angular instantânea do VCO em função da tensão de entrada.

³ Note-se que, dada a dependência do tempo, em rigor deveríamos escrever $\hat{\theta}(t)$ em vez de $\hat{\theta}$, o que às vezes faremos.

⁴ Tenha-se em conta que $\cos a \cos b = \frac{1}{2} [\cos(a+b) + \cos(a-b)]$ e $\sin a \cos b = \frac{1}{2} [\sin(a+b) + \sin(a-b)]$.

onde o índice k representa a iteração de ordem k . Seja em tempo contínuo, seja em tempo discreto, vemos que uma tensão de entrada do VCO positiva faz aumentar a fase da sinusóide de saída.

Substituindo valores na equação diferencial ficamos com

$$\frac{d\hat{\theta}(t)}{dt} = \frac{K_v K_l K_m}{2} [\theta - \hat{\theta}(t)] = K [\theta - \hat{\theta}(t)],$$

com $K = \frac{K_v K_l K_m}{2}$, em s^{-1} . A solução desta equação é

$$\hat{\theta}(t) = \theta + [\hat{\theta}(0) - \theta] e^{-Kt}.$$

A Fig. 17 mostra que, de facto, o PLL produz na saída do VCO uma sinusóide cuja fase, com o tempo, se ajusta e se aproxima da fase do sinal de entrada com uma velocidade de convergência que depende da constante K . Se, em vez do que acontece na Fig. 17, a estimativa inicial da fase for inferior ao valor verdadeiro, θ , a curva de aproximação é crescente.

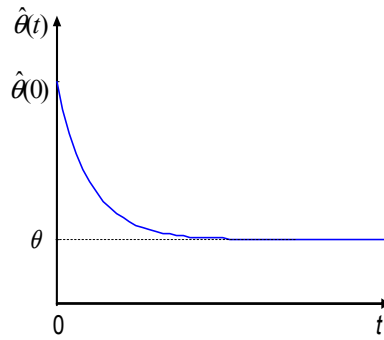


Fig. 17 Curva de aprendizagem do PLL.

Diz-se que no modo de seguimento o PLL está “locked”. Existe uma gama de valores de fase à volta de θ para a qual o PLL está “locked”. Fora desta gama o PLL não converge para o valor correcto.

Uma forma alternativa interessante de obter o erro de fase $\Delta\theta = \theta - \hat{\theta}$ entre um sinal $s(t) = \cos(2\pi f_c t + \theta)$ e uma sinusóide local $\hat{s}(t) = \cos(2\pi f_c t + \hat{\theta})$ é amostrar $s(t)$ nos instantes t_k em que $\hat{s}(t)$ é nula no sentido ascendente [3]. Nesses instantes especiais $\hat{s}(t_k) = \cos(2\pi f_c t_k + \hat{\theta}) = 0$ pelo que $2\pi f_c t_k + \hat{\theta} = -\pi/2$ e

$$2\pi f_c t_k + \theta = \underbrace{(2\pi f_c t_k + \hat{\theta})}_{-\pi/2} + \underbrace{(\theta - \hat{\theta})}_{\Delta\theta} = \Delta\theta - \pi/2.$$

As amostras do sinal de entrada valem, portanto,

$$s(t_k) = \cos(2\pi f_c t_k + \theta) = \cos(\Delta\theta - \pi/2) = \text{sen } \Delta\theta \approx \Delta\theta,$$

se $\Delta\theta$ for muito pequeno, demonstrando que à saída do detector obtemos realmente o erro de fase. A Fig. 18a ilustra este detector de fase por amostragem e a Fig. 18b exemplifica-o.

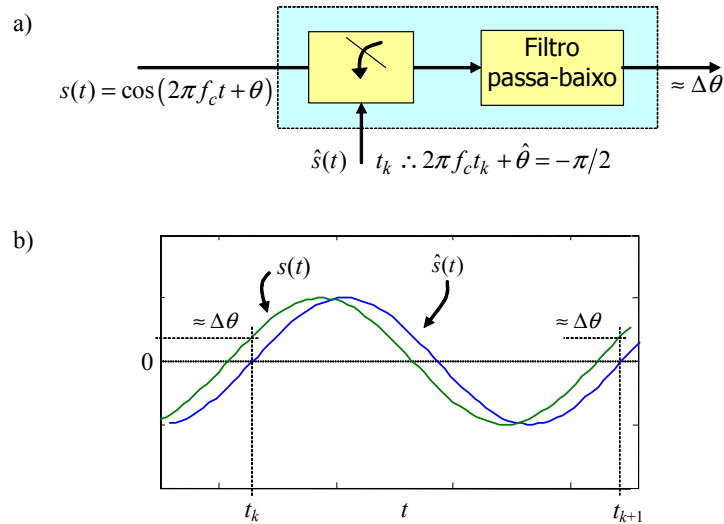


Fig. 18 O detector de fase por amostragem. a) diagrama de blocos; b) o erro de fase.

Vejamos agora como usar o PLL na recuperação da fase da portadora PSK.

3.1.2 Recuperação de fase com o “Costas loop”

Em 1956 John Costas propôs o circuito de recuperação de fase (desde então conhecido como “Costas Loop”) apresentado na Fig. 19. Suponhamos que não existe ruído e que é recebido o sinal BPSK $r(t) = Ab(t) \cos(2\pi f_c t + \theta)$, em que $b(t) = \pm 1$. À entrada do filtro passa-baixo do ramo em fase teremos

$$Ab(t) \cos(2\pi f_c t + \theta) \cos(2\pi f_c t + \hat{\theta}) = \frac{Ab(t)}{2} [\cos(2\pi \cdot 2f_c t + \theta + \hat{\theta}) + \cos(\theta - \hat{\theta})].$$

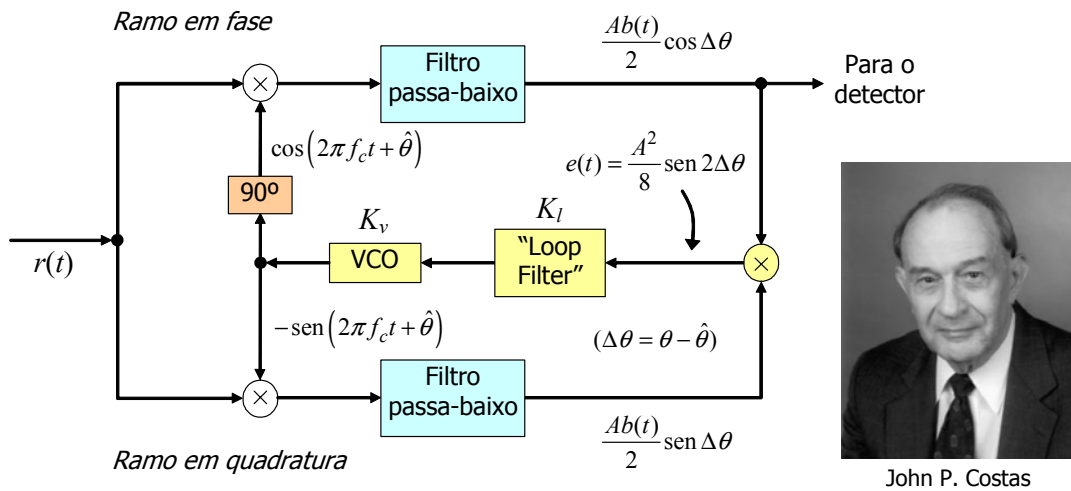


Fig. 19 “Costas loop”.

O filtro passa-baixo, de ganho unitário, elimina a componente em $2f_c$ pelo que na sua saída obteremos

$$\frac{Ab(t)}{2} \cos \Delta \theta .$$

Efectuando os mesmo cálculos para o ramo em quadratura encontramos à entrada do respectivo filtro

$$\begin{aligned}
-r(t)\sin(2\pi f_c t + \hat{\theta}) &= -Ab(t)\cos(2\pi f_c t + \theta)\sin(2\pi f_c t + \hat{\theta}) = \\
&= \frac{Ab(t)}{2} \left[\sin(\theta - \hat{\theta}) - \sin(2\pi \cdot 2f_c t + \theta + \hat{\theta}) \right]
\end{aligned}$$

e na sua saída

$$\frac{Ab(t)}{2} \sin \Delta\theta.$$

Tendo em conta que $b^2(t) = 1$ o “erro” $e(t)$, obtido através da multiplicação no PLL, vale

$$e(t) = \frac{A^2}{8} \sin 2\Delta\theta$$

Vemos assim que o funcionamento do circuito não depende da sequência de informação $b(t)$.

Sabemos que $\sin x \approx x$ se $x \ll 1$. Ora, como se disse, na prática o erro de estimação $\Delta\theta$ é muito pequeno pelo que, então,

$$\begin{aligned}
e(t) &\approx \frac{A^2}{4} \Delta\theta \\
e'(t) &= K_I e(t) \approx \frac{A^2 K_I}{4} \Delta\theta,
\end{aligned}$$

Na Sec. 3.1.1 encontrámos expressões que regem o funcionamento do PLL. Aplicando-as de novo e substituindo valores (com K_m unitário) obtemos:

$$\begin{aligned}
\frac{d\hat{\theta}(t)}{dt} &= K_v e'(t) \\
\frac{d\hat{\theta}(t)}{dt} &= \frac{A^2 K_I K_v}{4} [\theta - \hat{\theta}(t)] = K [\theta - \hat{\theta}(t)] \quad (\text{com } K = \frac{A^2 K_I K_v}{4}) \\
\hat{\theta}(t) &= \theta + [\hat{\theta}(0) - \theta] e^{-Kt}
\end{aligned}$$

Portanto, tal como no PLL simples, o VCO do “Costas Loop” gera uma sinusóide cuja fase, em regime permanente, é a estimativa da fase da portadora, $\hat{\theta} \approx \theta$, como se pretendia. É de notar, contudo, que esta estimativa é obtida quer a entrada seja $r(t) = Ab(t)\cos(2\pi f_c t + \theta)$ ou $r(t) = Ab(t)\cos(2\pi f_c t + \theta \pm \pi)$. Esta ambiguidade de fase de $\pm\pi$ rad pode ser eliminada através de codificação diferencial, como se sabe, ou então usando um procedimento “Data-Aided” com uma sequência de treino prévia na fase inicial da transmissão.

No modo de seguimento o erro de estimação é muito baixo ($\Delta\theta \approx 0$). Assim, no ramo em fase do “Costas loop” vamos ter um sinal $\frac{Ab(t)}{2} \cos \Delta\theta \approx \frac{Ab(t)}{2}$ proporcional à sequência de informação $b(t)$. O “Costas loop” pode, portanto, servir também para desmodular o sinal BPSK, como se indica na Fig. 19. Nesse sentido os filtros passa-baixo dos ramos em fase e em quadratura devem ser substituídos, com vantagem, por filtros adaptados.

3.1.3 Recuperação de fase com dispositivo não linear (multiplicador de frequência)

Como o espectro de um sinal MPSK não contém risca espectral à frequência da portadora, um modo de recuperar a fase desta é fazer passar o sinal recebido por um dispositivo não linear que crie uma risca a uma frequência múltipla de f_c , por exemplo, Mf_c . A portadora pretendida obtém-se por posterior divisão de frequência por M . Em PSK binário toma-se $M = 2$ (usando um dispositivo quadrático, como na Fig. 20) e em

QPSK toma-se $M = 4$. De um modo geral, em MPSK usa-se um circuito que eleve o sinal recebido à M -ésima potência. Vejamos o que se passa em BPSK.

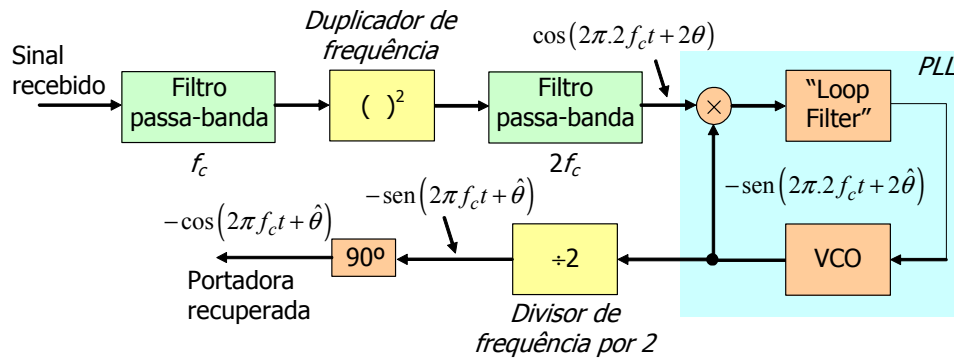


Fig. 20 Recuperação da fase da portadora BPSK com dispositivo quadrático.

Tendo à entrada do circuito o sinal $r(t) = Ab(t)\cos(2\pi f_c t + \theta)$ e não esquecendo que $b^2(t) = 1$, à saída do elemento quadrático obtemos uma componente contínua e uma risca à frequência $2f_c$:

$$r^2(t) = A^2 b^2(t) \cos^2(2\pi f_c t + \theta) = \frac{A^2}{2} [1 + \cos(2\pi \cdot 2f_c t + 2\theta)].$$

A componente contínua é eliminada pelo filtro sintonizado em $2f_c$. Com uma adequada normalização de amplitude o sinal à entrada do PLL é $\cos(2\pi \cdot 2f_c t + 2\theta)$, obrigando o VCO do PLL a gerar a sinusóide $-\sin(2\pi \cdot 2f_c t + 2\hat{\theta})$ da qual, por divisão de frequência por 2, se obtém a desejada estimativa $\sin(2\pi f_c t + \hat{\theta})$ – ou a simétrica, por haver ambiguidade de fase, como veremos já. Preferindo, pode também obter-se $\cos(2\pi f_c t + \hat{\theta})$.

Tal como o “Costas loop” este método, de facto, cria uma ambiguidade de fase de $\pm\pi$ radianos. Na verdade, se na entrada tivéssemos $Ab(t)\cos(2\pi f_c t + \theta \pm \pi) = -Ab(t)\cos(2\pi f_c t + \theta)$ em vez de $Ab(t)\cos(2\pi f_c t + \theta)$ a estimativa de fase seria a mesma. Para efeitos de sincronização de fase (inferior a π rad) é indiferente, portanto, obter uma sinusóide ou a sua simétrica à saída do VCO. O problema da ambiguidade resolve-se de novo com codificação diferencial ou com uma sequência de treino.

A Fig. 21 apresenta o circuito não linear a usar em MPSK. Desta vez a ambiguidade de fase é de $2\pi/M$ rad.

Para concluir, é de notar que a técnica de recuperação de fase com dispositivo não linear efectua o processamento na banda do canal, ao invés do “Costas loop” onde, após remoção da portadora, o PLL funciona em banda-base.

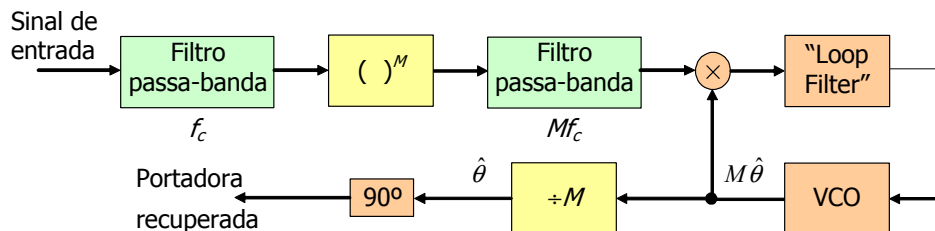


Fig. 21 Recuperação da fase da portadora MPSK com elevação à M -ésima potência

3.2 Sincronização de fase com malha aberta e estimação ML

Antes de abordarmos a sincronização por estimação ML vale a pena recordar, mesmo que de forma muito breve, os critérios de estimação MAP e ML.

3.2.1 Estimação MAP e ML

Suponhamos que desejamos obter a estimativa \hat{s} de um símbolo ou parâmetro binário s_i , $i = 0,1$, a partir da observação de um sinal z que de alguma forma está correlacionado com s_i . Sendo s_i binário só temos duas hipóteses de escolha: ou o sinal recebido z é devido a s_0 (hipótese H_0) ou é devido a s_1 (hipótese H_1).

O critério de máxima probabilidade a posteriori (MAP) estabelece que

$$p(s_1 | z) \underset{H_0}{\overset{H_1}{\geq}} p(s_0 | z), \quad (\text{MAP})$$

onde $p(s_i | z)$ representa a função densidade de probabilidade condicional de s_i dado que conhecemos z (probabilidade a posteriori). A expressão anterior significa que se $p(s_1 | z) > p(s_0 | z)$ devemos escolher a hipótese H_1 (ou seja, estimar $\hat{s}_{MAP} = s_1$); no caso contrário deveremos escolher a hipótese H_0 ($\hat{s}_{MAP} = s_0$).

O teorema de Bayes permite-nos escrever

$$p(s_1 | z) = \frac{p(z | s_1)p(s_1)}{p(z)} \underset{H_0}{\overset{H_1}{\geq}} \frac{p(z | s_0)p(s_0)}{p(z)} \quad \Rightarrow \quad \frac{p(z | s_1)}{p(z | s_0)} \underset{H_0}{\overset{H_1}{\geq}} \frac{p(s_0)}{p(s_1)}$$

A função densidade de probabilidade condicional $p(z|s_i)$ é designada por *função de verosimilhança* de s_i .

Se os símbolos s_0 e s_1 forem equiprováveis então $p(s_0) = p(s_1)$ e obtemos o critério de máxima verosimilhança, ou critério ML:

$$p(z | s_1) \underset{H_0}{\overset{H_1}{\geq}} p(z | s_0). \quad (\text{ML})$$

As expressões anteriores não calculam um valor explícito de \hat{s}_{MAP} ou de \hat{s}_{ML} mas são equivalentes às seguintes formulações alternativas:

$$\text{Critério MAP: } \hat{s}_{MAP} = \arg \max_{s_i} p(s_i | z)$$

$$\text{Critério ML: } \hat{s}_{ML} = \arg \max_{s_i} p(z | s_i)$$

Nota: “arg max” indica o argumento do máximo da função. Se, por exemplo, uma função $f(x)$ atinge o máximo para $x = a$ então $\max_x f(x) = f(a)$ e $\arg \max_x f(x) = a$.

Em resumo:

- Na estimação MAP consideramos probabilidades a posteriori $p(s_i | z)$, isto é, probabilidades condicionais obtidas após observação do sinal recebido. Os símbolos ou parâmetros s_i são considerados aleatórios.
- Na estimação ML consideramos probabilidades a priori $p(z|s_i)$ (verosimilhanças), isto é, probabilidades condicionais já conhecidas antecipadamente. Os símbolos ou parâmetros s_i são considerados determinísticos, embora desconhecidos.
- Se os símbolos forem equiprováveis os dois critérios de estimação são equivalentes ($\hat{s}_{MAP} = \hat{s}_{ML}$).

Vamos restringir-nos à estimação ML pois estamos a considerar que os símbolos são equiprováveis.

3.2.2 Recuperação de fase em tempo contínuo através de estimação ML

Na estimação ML de um parâmetro de sinal, como θ , é necessário, portanto, determinar a função densidade de probabilidade do sinal recebido condicionada ao valor desse parâmetro, o que requer que o sinal recebido seja observado no receptor durante um longo intervalo de tempo $T_0 \gg T$. A este intervalo T_0 chama-se, naturalmente, *intervalo de observação*.

Vamos supor que o sinal BPSK $s(t)$ é recebido com ruído gaussiano branco aditivo $n(t)$ de densidade espectral de potência $N_0/2$:

$$\begin{aligned} r(t) &= s(t) + n(t) = \\ &= Ab(t)\cos(2\pi f_c t + \theta) + n(t) \end{aligned}$$

Devido a θ necessitamos de um espaço de sinal de duas dimensões. Podendo as funções-base que o definem⁵ ser $\psi_1(t) = \cos 2\pi f_c t$ e $\psi_2(t) = \sin 2\pi f_c t$, à saída do correlacionador teremos as componentes

$$\begin{aligned} z_1 &= \int_{T_0} r(t)\psi_1(t)dt = Ab(t)\int_{T_0} \cos(2\pi f_c t + \theta)\cos 2\pi f_c t dt + \int_{T_0} n(t)\psi_1(t)dt = \\ &= \underbrace{\frac{A}{2}b(t)T_0 \cos \theta}_{s_1} + n_1 \\ z_2 &= \int_{T_0} r(t)\psi_2(t)dt = -\underbrace{\frac{A}{2}b(t)T_0 \sin \theta}_{s_2} + n_2 \end{aligned}$$

ou

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} s_1 \\ s_2 \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} = \begin{bmatrix} \frac{A}{2}b(t)T_0 \cos \theta \\ -\frac{A}{2}b(t)T_0 \sin \theta \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \end{bmatrix}.$$

A função de verosimilhança que pretendemos obter para procedermos à estimação ML é a função densidade de probabilidade conjunta $p(z_1, z_2 | \theta)$. Vamos ver como.

Prova-se que as componentes de ruído n_1 e n_2 são variáveis aleatórias gaussianas independentes, de média nula e variância $\sigma^2 = N_0/2$, de funções densidade de probabilidade

$$p(n_1) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{n_1^2}{N_0}\right) \quad p(n_2) = \frac{1}{\sqrt{\pi N_0}} \exp\left(-\frac{n_2^2}{N_0}\right)$$

Por conseguinte, as componentes z_1 e z_2 são variáveis aleatórias gaussianas de variância σ^2 e média s_1 e s_2 , respectivamente. Como são independentes, a função densidade de probabilidade conjunta $p(z_1, z_2 | \theta)$ é igual ao produto das funções densidade de probabilidade individuais:

$$\begin{aligned} p(z_1, z_2 | \theta) &= p(z_1 | \theta)p(z_2 | \theta) = \\ &= \frac{1}{\pi N_0} \exp\left(-\frac{(z_1 - s_1)^2 + (z_2 - s_2)^2}{N_0}\right) \end{aligned}$$

⁵ Os resultados a que vamos chegar são os mesmos quer o espaço de sinal seja ou não ortonormado. Sem perda de generalidade, optou-se por funções-base de energia não-unitária apenas para simplificar expressões.

É esta a função de verosimilhança pretendida. O estimador ML de fase é o valor de θ para o qual $p(z_1, z_2 | \theta)$ atinge o seu máximo:

$$\hat{\theta}_{ML} = \arg \max_{\theta} p(z_1, z_2 | \theta).$$

Mas se $\hat{\theta}_{ML}$ maximiza $p(z_1, z_2 | \theta)$ então claramente também maximiza o logaritmo natural dessa função:

$$\hat{\theta}_{ML} = \arg \max_{\theta} [\ln p(z_1, z_2 | \theta)].$$

Esta forma alternativa tem a vantagem de tornar as expressões e os cálculos mais simples. Ora

$$\ln p(z_1, z_2 | \theta) = \ln \frac{1}{\pi N_0} - \frac{(z_1 - s_1)^2 + (z_2 - s_2)^2}{N_0}.$$

Tendo em consideração apenas os termos que dependem de θ (apenas $z_1 s_1$ e $z_2 s_2$) conclui-se, após algum desenvolvimento, que

$$\hat{\theta}_{ML} = \arg \max_{\theta} (z_1 \cos \theta - z_2 \sin \theta)$$

O máximo da função $z_1 \cos \theta - z_2 \sin \theta$ ocorre quando a sua derivada é nula,

$$\frac{d}{d\theta} (z_1 \cos \theta - z_2 \sin \theta) = -z_1 \sin \hat{\theta}_{ML} - z_2 \cos \hat{\theta}_{ML} = 0,$$

donde se tira, finalmente,

$$\begin{aligned} \hat{\theta}_{ML} &= \arctan \left(-\frac{z_2}{z_1} \right) = \\ &= \arctan \left(-\frac{\int_{T_0} r(t) \sin 2\pi f_c t dt}{\int_{T_0} r(t) \cos 2\pi f_c t dt} \right) \end{aligned}$$

A Fig. 22 apresenta um esquema de cálculo do estimador de fase ML.

A estimativa ML da fase da portadora também pode ser obtida iterativamente através de um algoritmo semelhante ao algoritmo LMS do processamento adaptativo de sinais [4].

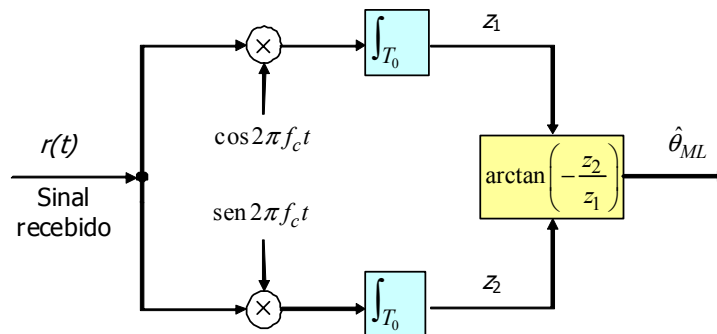


Fig. 22 Esquema de cálculo do estimador de fase de máxima verosimilhança.

4. Recuperação da fase da portadora em tempo discreto

Consideremos agora que todo o processamento é feito com base em amostras recolhidas nos instantes de amostragem adequados, ou seja, conhecemos a temporização dos símbolos, como antes, mas deixamos de trabalhar em tempo contínuo para passar para tempo amostrado. Suponhamos então que à saída do correlacionador ou do filtro adaptado dispomos de amostras complexas de sinal e ruído $z(1), z(2), \dots, z(k), \dots$ – complexos $z(k)$ cujas partes real e imaginária são, respectivamente, a componente em fase $z_1(k)$ e a componente em quadratura $z_2(k)$. De novo várias abordagens são possíveis: ou usamos os dados (método DA), ou as decisões do receptor (método DD) ou simplesmente não usamos auxílios (método NDA). Em qualquer caso iremos obter estimativas de máxima verosimilhança.

Começemos pelos métodos ML com malha aberta (DA e NDA).

4.1.1 Recuperação de fase em tempo discreto com a ajuda dos dados (ML DA)

Suponhamos que durante um dado intervalo de observação temos à saída do correlacionador ou do filtro adaptado L amostras complexas de sinal e ruído $z(k)$, para $k=1, 2, \dots, L$. Se também conhecermos igual número de símbolos enviados c_k a estimativa ML DA da fase da portadora é dada pela expressão simples

$$\hat{\theta} = \arg \left\{ \sum_{k=1}^L c_k^* z(k) \right\},$$

na qual $\arg(\mathbf{x}) = \arctg \left[\frac{\text{Im}(\mathbf{x})}{\text{Re}(\mathbf{x})} \right]$ é a fase (ângulo) do complexo \mathbf{x} na sua representação polar [5].

Corresponde-lhe o diagrama de blocos da Fig. 23.

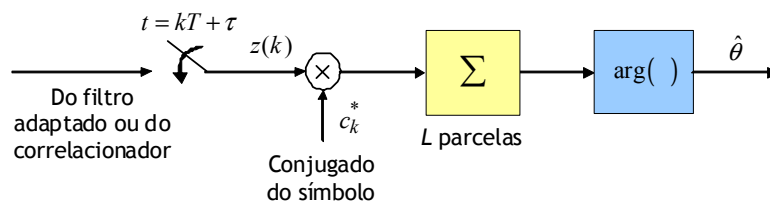


Fig. 23 Recuperação ML da fase com temporização conhecida e ajuda dos símbolos.

4.1.2 Recuperação de fase em tempo discreto sem a ajuda dos dados (ML NDA)

A recuperação da fase da portadora em modulações cuja constelação tem geometria rotacionalmente simétrica, como em MPSK e QAM, pode ser realizada sem a ajuda dos dados combinando o dispositivo não linear da Sec. 3.1.3 com a estimação de máxima verosimilhança da Sec. 3.2.2 [5] [6]. O método tem a vantagem de dispensar o PLL. Assim, em constelações invariantes à rotação de $2\pi/N$ a estimativa NDA da fase da portadora é calculada de acordo com o diagrama da Fig. 24 e seguindo as fórmulas

$$\text{MPSK: } \hat{\theta} = \frac{1}{M} \arg \left[\sum_{k=1}^L z^M(k) \right] \quad \text{QAM: } \hat{\theta} = \frac{1}{4} \arg \left[-\sum_{k=1}^L z^4(k) \right]$$

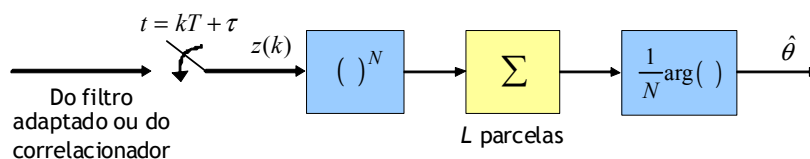


Fig. 24 Recuperação ML da fase com temporização conhecida mas sem ajuda dos símbolos

As constelações MPSK e QAM são invariantes a rotações de $2\pi/M$ e $\pi/2$, respectivamente, ou seja, $N=M$ em MPSK e $N=4$ em QAM, independentemente da constelação, mostrando que a potência a usar nestes estimadores de fase é sempre N .

Se $z(k) = |z(k)|e^{j\phi(k)}$ for a representação polar da amostra complexa $z(k)$ então a expressão da estimativa da fase em MPSK escreve-se $\hat{\theta} = \frac{1}{M} \arg \left\{ \sum_{k=1}^L |z(k)|^M e^{jM\phi(k)} \right\}$. Uma generalização desta equação, proposta em [7] para sinais MPSK, obtém-se substituindo o termo $|z(k)|^M$ por uma função $F(|z(k)|)$ adequada:

$$\hat{\theta} = \frac{1}{M} \arg \left\{ \sum_{k=1}^L F(|z(k)|) e^{jM\phi(k)} \right\} \quad (MPSK).$$

Se $F(|z(k)|) = |z(k)|^M$ regressamos à expressão original. Têm-se conseguido melhores desempenhos com potências inferiores a M , por exemplo com $F(|z(k)|) = |z(k)|^2$ ou mesmo $F(|z(k)|) = 1$, dependendo da relação sinal-ruído e do número de pontos da constelação [5].

A estimação de fase com a função $F(|z(k)|)$ é normalmente designada pelo nome dos autores: *algoritmo de Viterbi & Viterbi*.

4.1.2.1 Recuperação de fase em MPSK sem se conhecer a temporização de símbolos

Temos suposto que a temporização de símbolo é conhecida. Caso tal não aconteça o algoritmo de estimação ML precedente pode ainda ser utilizado em MPSK desde que a amostragem do sinal à saída do correlacionador ou do filtro adaptado seja realizada a uma cadência mais elevada [5]: em vez de símbolo a símbolo (de T em T segundos) deve ser feita várias vezes por símbolo (digamos, N_s vezes, nos instantes $t = kT/N_s$). A nova estimativa passa a basear-se em $N_s L$ amostras $z(kT/N_s)$, recolhidas de T/N_s em T/N_s segundos, e a expressão respectiva é

$$\text{MPSK:} \quad \hat{\theta} = \frac{1}{M} \arg \left[\sum_{k=1}^{N_s L} z^M(kT/N_s) \right]$$

Suponhamos que usamos impulsos de cosseno elevado com factor de *roll-off* α . Qual deve ser o valor do factor de sobreamostragem N_s ? Sabemos que a largura de banda dos impulsos de cosseno elevado é igual a $\frac{1}{2T}(1+\alpha)$. Ora a elevação à potência M alarga a banda ocupada M vezes, passando para $\frac{M}{2T}(1+\alpha)$. De acordo com o teorema da amostragem a frequência de amostragem deve ser, no mínimo, o dobro da maior frequência do sinal, ou seja, $f_s \geq 2 \frac{M}{2T}(1+\alpha)$. Mas $f_s = N_s/T$. Portanto,

$$N_s \geq M(1+\alpha).$$

4.1.3 Recuperação de fase em tempo discreto com a ajuda das decisões (ML DD)

O método com a ajuda das decisões é o único dos métodos em tempo discreto apresentados que usa realimentação, isto é, malha fechada, correspondendo-lhe o diagrama de blocos da Fig. 25 [5].

Ao contrário dos métodos DA e NDA, neste não se recorre a um intervalo de observação; pelo contrário, o método vai calculando iterativamente uma estimativa da fase, instante a instante melhorando-a, de acordo com o seguinte algoritmo:

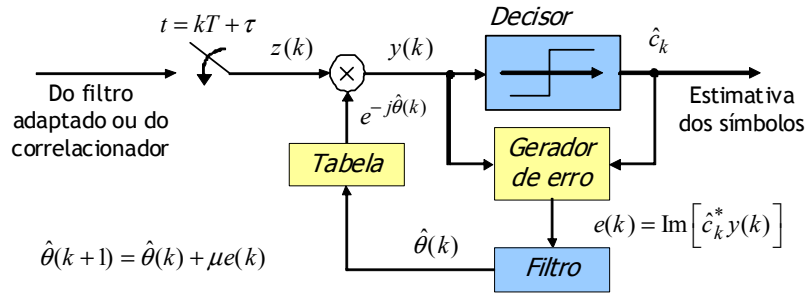


Fig. 25 Esquema de recuperação de fase ML DD.

1. $y(k) = z(k)e^{-j\hat{\theta}(k)}$
2. $e(k) = \text{Im}[\hat{c}_k^* y(k)]$
3. $\hat{\theta}(k+1) = \hat{\theta}(k) + \mu e(k)$

O complexo \hat{c}_k representa a estimativa do ponto enviado c_k e o coeficiente μ é uma pequena constante positiva (inferior a 1) chamada *passo de adaptação*.

O bloco *Gerador de Erro* produz na sua saída aproximadamente o erro de fase $e(k) = \theta - \hat{\theta}(k)$, como se mostra a seguir. Uma das entradas do bloco, $y(k) = z(k)e^{-j\hat{\theta}(k)}$, é o ponto recebido rodado do ângulo de fase estimado nesse momento e a outra, \hat{c}_k , é o símbolo estimado. Suponhamos, para simplificar, que não existe ruído e que o ponto enviado, c_k , sofreu apenas uma rotação de θ radianos: $z(k) = c_k e^{j\theta}$. Assim,

$$\hat{c}_k^* y(k) = \hat{c}_k^* z(k) e^{-j\hat{\theta}(k)} = \hat{c}_k^* c_k e^{j(\theta - \hat{\theta}(k))}$$

Se a decisão \hat{c}_k estiver correcta então $\hat{c}_k = c_k$ e $\hat{c}_k^* c_k = c_k^* c_k = |c_k|^2 = 1$ (em MPSK, claro!). Logo,

$$\hat{c}_k^* y(k) = e^{j(\theta - \hat{\theta}(k))} = \cos(\theta - \hat{\theta}(k)) + j \text{sen}(\theta - \hat{\theta}(k))$$

Em regime permanente $\hat{\theta}(k) \approx \theta$, isto é, $\theta - \hat{\theta}(k) \ll 1$, situação em que $\cos(\theta - \hat{\theta}(k)) \approx 1$ e $\text{sen}(\theta - \hat{\theta}(k)) \approx \theta - \hat{\theta}(k)$ e, portanto, $\hat{c}_k^* y(k) \approx 1 + j(\theta - \hat{\theta}(k))$. Vê-se bem que a parte imaginária de $\hat{c}_k^* y(k)$ é o erro de fase $e(k) = \theta - \hat{\theta}(k)$, daí que $e(k) = \text{Im}[\hat{c}_k^* y(k)]$.

Exemplo 3:

Aplicou-se o método ML DD à recuperação da fase de uma portadora 8-PSK que chegou ao receptor com um esfasamento de $\theta = 12^\circ$. Considerou-se ruído gaussiano aditivo de variância 0,01, um valor inicial da estimativa da fase de 0° e um passo de adaptação $\mu = 0,025$. Na Fig. 26 pode observar-se a evolução da estimativa da fase ao longo de 500 iterações em direcção a 12° , valor assinalado na figura pela linha horizontal.

A Fig. 27 mostra a representação no espaço de sinal dos pontos recebidos: à esquerda nota-se que a constelação está rodada de um certo ângulo (que nós sabemos ser 12°) e à direita observa-se que essa rotação foi compensada (através do último valor de $\hat{\theta}(k)$, $\hat{\theta}(500) = 12,034^\circ$). Olhe-se para os eixos para se notar a diferença.

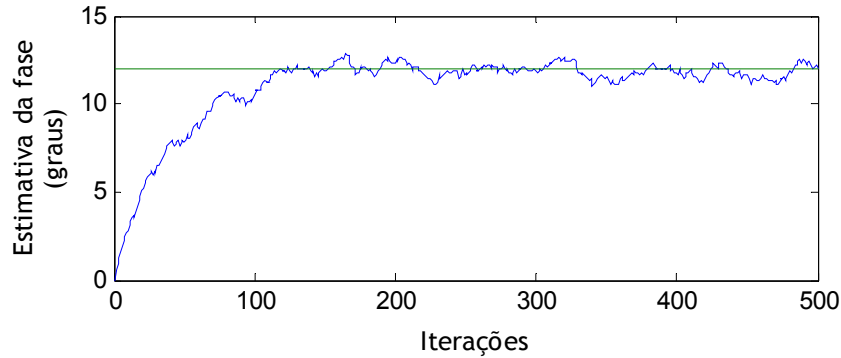


Fig. 26 Curva de aprendizagem da estimativa ML DD da fase da portadora.

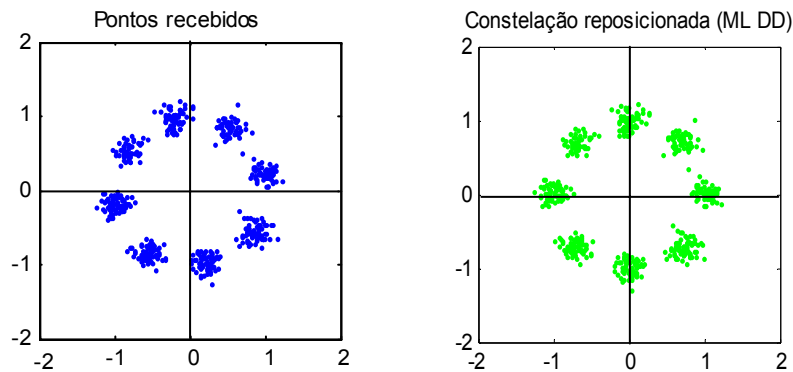


Fig. 27 Representação no espaço de sinal dos pontos 8-PSK. À esquerda: pontos recebidos; à direita: pontos reposicionados após recuperação da fase da portadora.

5. Recuperação da temporização de símbolo

A sincronização de símbolo – ou seja, a estimação do atraso τ – pode ser realizada enviando um sinal de relógio (bits de sincronismo) juntamente com o sinal modulado. Esta opção “Data-Aided”, todavia, apresenta duas desvantagens: por um lado, parte da potência disponível é gasta no envio do sinal extra, o que é um desperdício; por outro, no receptor esse sinal extra tem de ser adequadamente separado do sinal modulado para que não interfiram um com o outro. A abordagem alternativa é recuperar o relógio exclusivamente a partir do sinal modulado recebido, opção que seguiremos.

Tal como na sincronização de fase da portadora, também a sincronização de símbolo pode ser efectuada com ou sem realimentação. No primeiro caso pode usar-se, por exemplo, um circuito com PLL designado por *sincronizador “Early-Late”* ou um sincronizador em tempo discreto baseado no chamado *método de Gardner*; no segundo caso (malha aberta) basta um simples circuito com dispositivo quadrático. A recuperação de relógio com malha aberta pode ainda realizar-se através de uma estimação de máxima verosimilhança. Matematicamente complicada, a estimação ML do atraso τ ultrapassa o nível de profundidade e dificuldade pretendido com este texto. Os interessados poderão consultar a bibliografia indicada, em particular as obras de Proakis e Haykin.

No que se segue, suporemos sempre que a fase da portadora foi recuperada de maneira correcta (logo, $\hat{\theta} = \theta$).

5.1 Sincronização de símbolo com dispositivo quadrático

A sincronização de símbolo com dispositivo não linear (usualmente de lei quadrática) é um método clássico de *tempo contínuo* (ao contrário de outros que, recorrendo a amostras, são métodos de *tempo discreto* ou *tempo amostrado*). Um diagrama de blocos muito usado adequado a PSK binário é apresentado na Fig. 28.

A entrada do circuito é uma sequência binária em banda-base, obtida após multiplicação do sinal modulado pela portadora de referência local $\cos 2\pi f_c t$. O primeiro filtro sintonizado⁶ ajuda a reduzir o ruído que acompanha a sequência de bits e a elevação ao quadrado (rectificação) cria uma componente espectral na frequência $1/T$, componente que é filtrada pelo filtro sintonizado de banda estreita centrada nessa frequência e de que resulta uma onda sinusoidal de temporização apropriada para definir os instantes de amostragem a usar na detecção dos bits de dados.

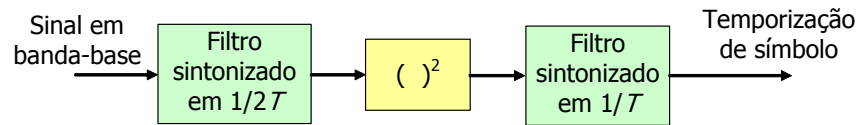


Fig. 28 Sincronizador de símbolos com dispositivo quadrático para sinais BPSK.

Com sinais bidimensionais (MPSK e QAM) poderá usar-se o diagrama da Fig. 29 [8]. Na entrada são aplicadas as componentes em fase e em quadratura, sequências de símbolos obtidos, como é costume, através da multiplicação prévia do sinal modulado pelas respectivas funções-base $\cos 2\pi f_c t$ e $\sin 2\pi f_c t$. Esses dois sinais em banda-base são então filtrados por filtros de banda estreita centrada em $1/2T$, elevados ao quadrado (para que se obtenham componentes espectrais no dobro da frequência), somados e depois, finalmente, filtrados pelo filtro de banda estreita sintonizado na frequência de relógio, $1/T$, obtendo-se assim a onda sinusoidal de temporização adequada à recuperação dos símbolos.

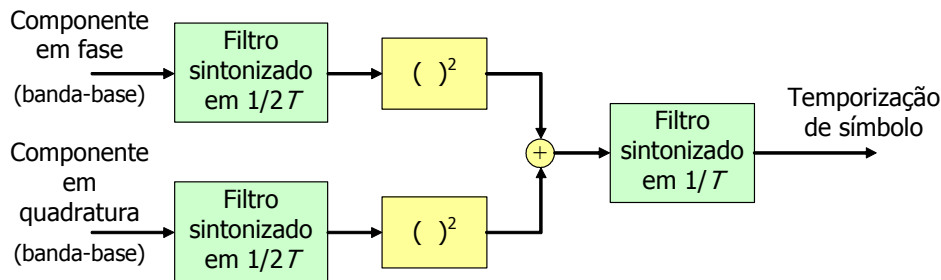


Fig. 29 Sincronizador de símbolos com dispositivo quadrático para sinais MPSK e QAM.

Exemplo 4:

Criou-se o modelo Simulink da Fig. 30 para exemplificar a recuperação de relógio através de dispositivo quadrático. Os dados apresentam-se na forma de uma sequência polar de impulsos rectangulares ± 1 à razão de $1/T = 1000$ bits/s e são aplicados a um filtro adaptado amostrado ("Discrete Filter") realizado como um filtro FIR. O filtro passa-banda de banda estreita ("BPF Filter") tem frequência central 1000 Hz.

Na Fig. 31 podem ver-se as formas de onda recolhidas em vários pontos do modelo. Em particular notamos a esperada forma de onda triangular/trapezoidal de saída do filtro adaptado ("Scope2"), a sua rectificação ("Scope3") e a quase sinusóide de saída do filtro final ("Scope4"). Pode observar-se na Fig. 32 que esta forma de onda tem uma fortíssima componente espectral em $1/T = 1000$ Hz. A variação de amplitude da envolvente visível em "Scope4" é eliminada pelo bloco "Sign" – cuja saída é ± 1 consoante o sinal da entrada – obtendo-se, finalmente, uma onda quadrada de período $T = 1$ ms ("Scope5").

⁶ Este primeiro filtro sintonizado é, por vezes, substituído por um filtro adaptado aos impulsos de formatação de bits [9].

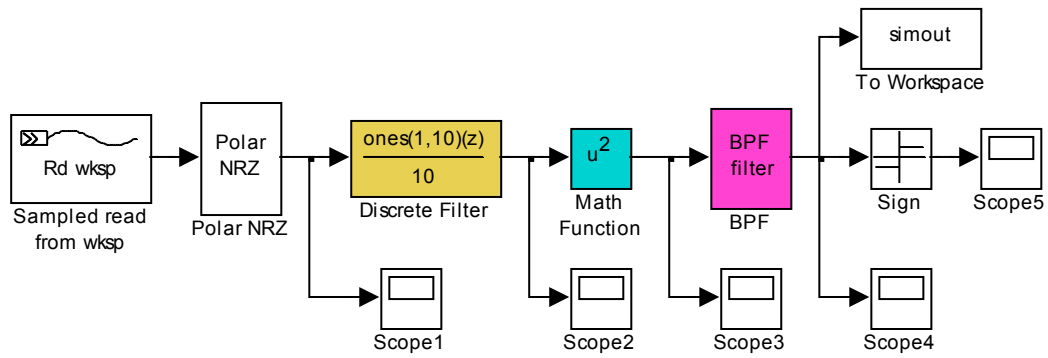
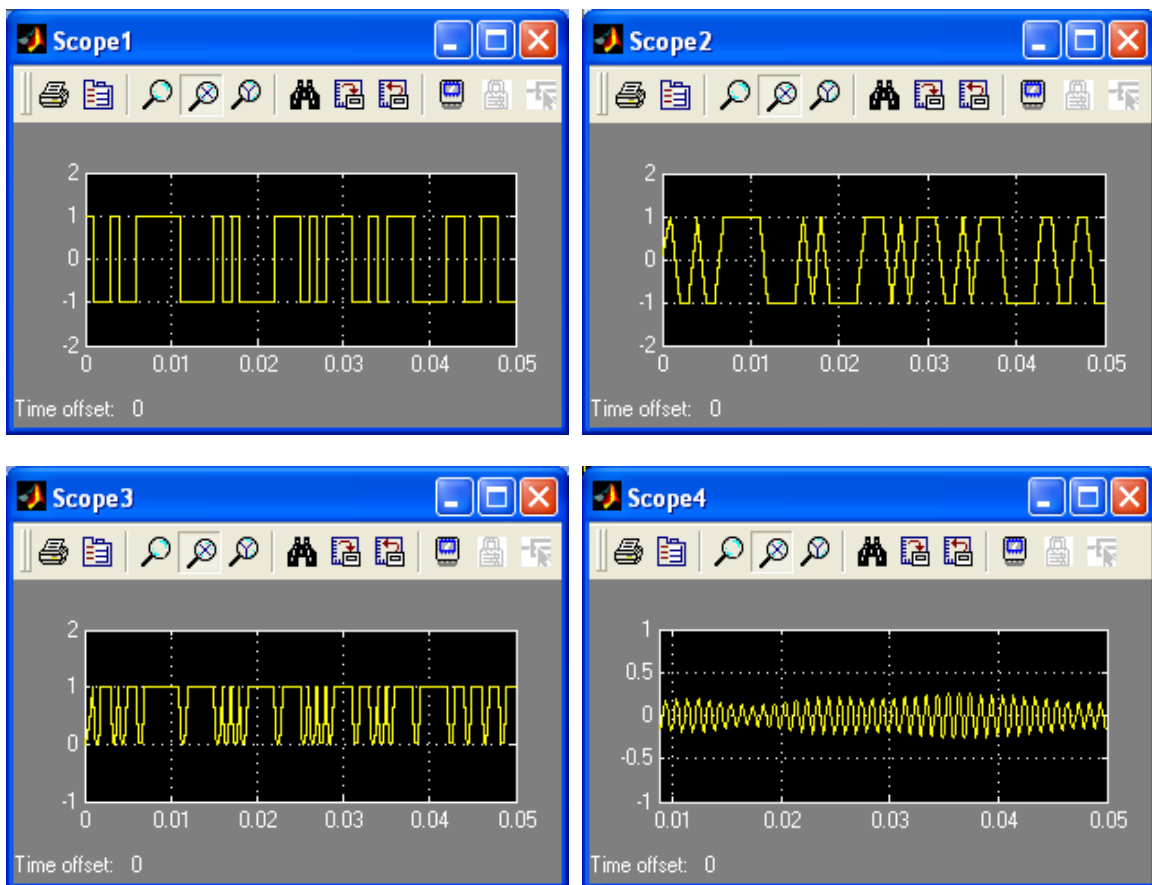


Fig. 30 Modelo Simulink simples para recuperação de relógio por dispositivo quadrático.



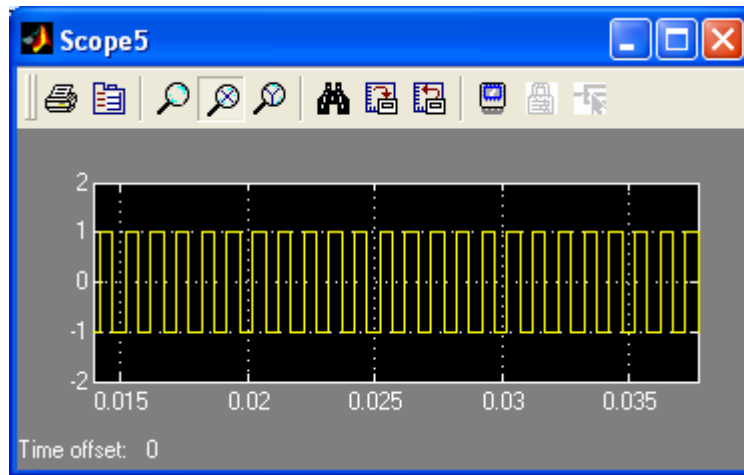


Fig. 31 Formas de onda obtidas em vários pontos do modelo da Fig. 30.

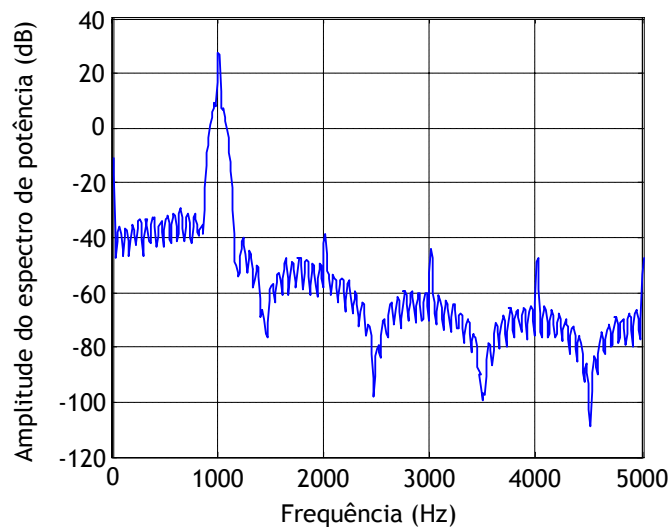


Fig. 32 Espectro da forma de onda do “Scope4” da Fig. 31.

5.2 O sincronizador “Early-Late”

O sincronizador “Early-Late” é um esquema muito vulgar de recuperação de temporização com realimentação. Diz-se que é um método de *tempo amostrado* pois, como se pode ver na Fig. 33, funciona com amostras de sinal. Da multiplicação do sinal modulado de entrada pela portadora previamente recuperada $\cos(2\pi f_c t + \hat{\theta})$ obtém-se uma sequência em banda-base de símbolos de duração T e cadência $1/T$ símbolos/s. Pretende-se que o VCO produza uma sinusóide de período T (ou frequência $1/T$) sincronizada com os símbolos (isto é, quando um símbolo começa a onda sinusoidal deve passar por zero), pelo que as sucessivas estimativas $\dots, \hat{\tau}_{k-1}, \hat{\tau}_k, \hat{\tau}_{k+1}, \dots$ se vão aproximando do valor correcto mas desconhecido τ . Vejamos então como o circuito funciona, admitindo impulsos de formatação rectangulares [9].

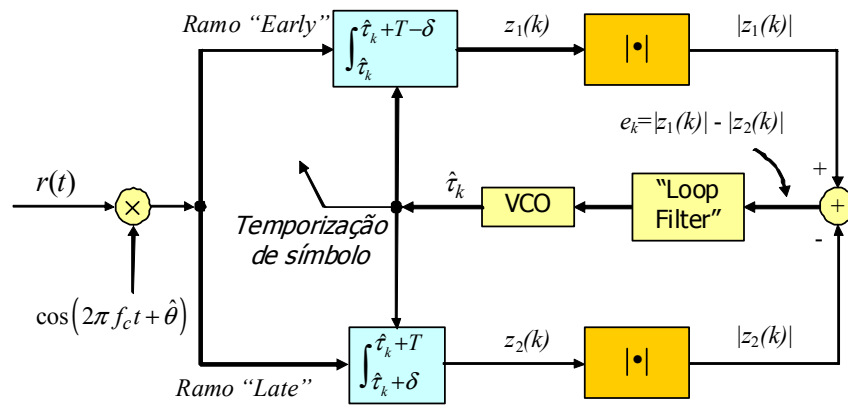


Fig. 33 O sincronizador “Early-Late”.

- A sequência em banda-base é aplicada a dois integradores com intervalo de integração de $T - \delta$ segundos, com $\delta < T/2$. No ramo de cima, designado por “Early” (adiantado), a integração começa no instante que se presume ser o instante de início do símbolo ($t = \hat{\tau}_k$) e acaba $T - \delta$ segundos depois, em $t = \hat{\tau}_k + T - \delta$, antes do símbolo terminar⁷ (mas o valor do integral é guardado até ao instante $t = \hat{\tau}_k + T$); no outro ramo, designado por “Late” (atrasado), a integração começa δ segundos após o início da primeira e termina em $t = \hat{\tau}_k + T$. Os rectificadores do sincronizador recebem depois os valores $z_1(k)$ e $z_2(k)$ provenientes dos integradores e as suas saídas geram o erro $e_k = |z_1(k)| - |z_2(k)|$, aplicado em seguida ao PLL.
- Quando a estimação do instante $t = \tau$ está correcta ($\hat{\tau}_k = \tau$, na situação A da Fig. 34), ambos os intervalos de integração estão contidos na duração T de um único símbolo de informação. Nesse caso os contributos $|z_1(k)|$ e $|z_2(k)|$ são iguais e o erro e_k é nulo, pelo que o VCO não altera a frequência da sua onda de saída;

$\hat{\tau}_k = \tau$ (relógio certo) $\Rightarrow e_k = |z_1(k)| - |z_2(k)| = 0 \Rightarrow$ frequência do VCO mantém-se $\Rightarrow \hat{\tau}_k$ não varia

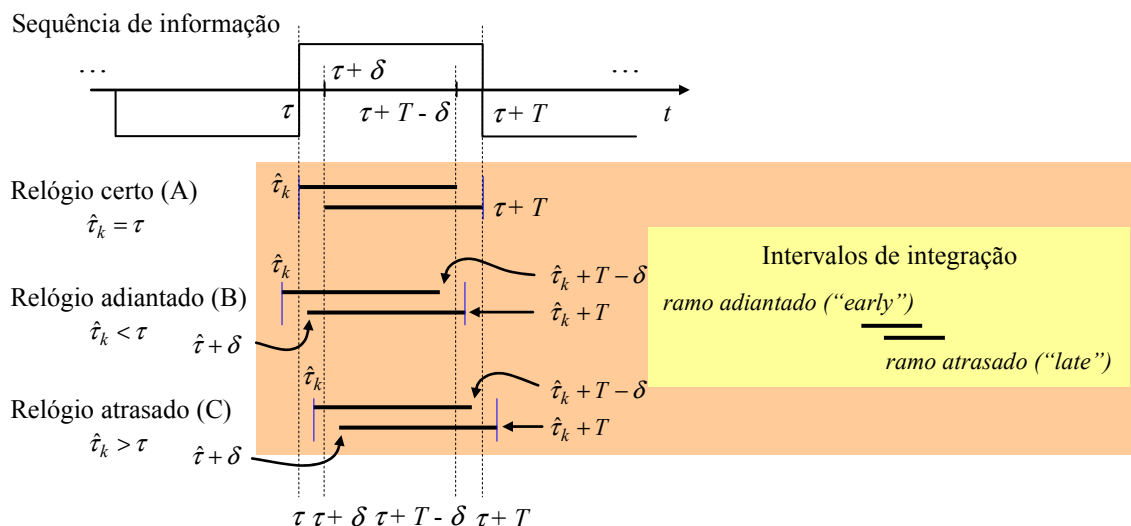


Fig. 34 Relógio certo, adiantado e atrasado no sincronizador “Early-Late”.

⁷ Em rigor o símbolo de ordem k começa no instante $t_{k-1} = (k-1)T + \tau$ e termina no instante $t_k = kT + \tau$. No entanto, para simplificar, suporemos que cada símbolo começa em $t = \tau$, sem perda de generalidade.

⁹ Na literatura de língua inglesa este bloco é designado por TED, de “Timing Error Detector”.

- Se, pelo contrário, a estimativa do instante $t = \tau$ estiver adiantada, isto é, se $\hat{t}_k < \tau$, como na situação B da Fig. 34, todo o intervalo de integração do ramo “late” está contido na duração de um símbolo ao passo que o outro intervalo se distribui por dois símbolos de polaridade oposta, o próprio e o anterior (o que diminui o valor absoluto, $|z_1(k)|$, do integral respectivo) – logo, $|z_2(k)| > |z_1(k)|$ e $e_k = |z_1(k)| - |z_2(k)| < 0$. Este erro negativo diminui a frequência do VCO e, com isso, retarda a estimativa do instante $t = \tau$, isto é, atrasa o relógio, como se pretende;

$$\hat{t}_k < \tau \text{ (relógio adiantado)} \Rightarrow e_k = |z_1(k)| - |z_2(k)| < 0 \Rightarrow \text{frequência do VCO baixa} \Rightarrow \hat{t}_k \text{ aumenta}$$

- Se, ao invés, o relógio local estiver atrasado ($\hat{t}_k > \tau$, na situação C da Fig. 34) só a integração do ramo “early” se faz durante um único símbolo pelo que $|z_1(k)| > |z_2(k)|$ e $e_k = |z_1(k)| - |z_2(k)| > 0$, obrigando o VCO a aumentar a frequência para acertar o relógio, adiantando-o.

$$\hat{t}_k > \tau \text{ (relógio atrasado)} \Rightarrow e_k = |z_1(k)| - |z_2(k)| > 0 \Rightarrow \text{frequência do VCO sobe} \Rightarrow \hat{t}_k \text{ diminui}$$

Note-se que considerámos impulsos consecutivos de polaridade oposta. Se os impulsos consecutivos tiverem a mesma polaridade as saídas dos ramos em fase e em quadratura são iguais, o VCO não altera a sua frequência e \hat{t}_k não é corrigido. Aqui se vê a conveniência (melhor dizendo, a necessidade) de garantir transições regulares por zero dos impulsos de dados para que a sincronização dos símbolos seja facilitada.

As considerações anteriores conduzem-nos à equação de actualização da temporização \hat{t} em instantes sucessivos

$$\hat{t}_{k+1} = \hat{t}_k - \mu T e_k,$$

em que $\mu > 0$ é o *passo de adaptação*.

O erro $e_k = |z_1(k)| - |z_2(k)|$ pode ser calculado sem atender aos valores absolutos, dispensando os rectificadores da Fig. 33. Seja $z(k)$ o integral de $t = \hat{t}_k$ a $t = \hat{t}_k + T$, isto é, seja $z(k)$ o valor a aplicar ao decisor. Este, em face do sinal de $z(k)$, estima o k -ésimo bit, $\hat{b}_k = \pm 1$. O erro pode então ser dado por

$$e_k = \hat{b}_k [z_1(k) - z_2(k)].$$

Com a ajuda da Fig. 36 façamos a verificação com os valores $z_1(k) = 3$, $z_2(k) = 2$, $z(k) = 5$ e os seus simétricos. Tenha-se em conta que se $z(k) = 5$ então $\hat{b}_k = +1$ e se $z(k) = -5$ então $\hat{b}_k = -1$. Os valores de e_k devem ser todos iguais, como iremos confirmar:

$$\begin{array}{lll} e_k = |z_1(k)| - |z_2(k)| : & e_k = |3| - |2| = 1 & e_k = |-3| - |-2| = 1 \\ e_k = \hat{b}_k [z_1(k) - z_2(k)] : & e_k = (+1) \times [3 - 2] = 1 & e_k = (-1) \times [(-3) - (-2)] = 1 \end{array}$$

Exemplo 5:

A Fig. 35 apresenta o diagrama de blocos de um sincronizador “Early-Late” associado a um filtro adaptado. É necessário recolher três amostras do sinal por cada símbolo, duas para o processo de sincronização e uma, a central, para a amostragem do sinal propriamente dita (no fundo, o objectivo de tudo). As três amostragens ocorrem nos instantes $t_k - \delta$, t_k e $t_k + \delta$, com $t_k = kT + \hat{t}_k$ e $\delta < T/2$.

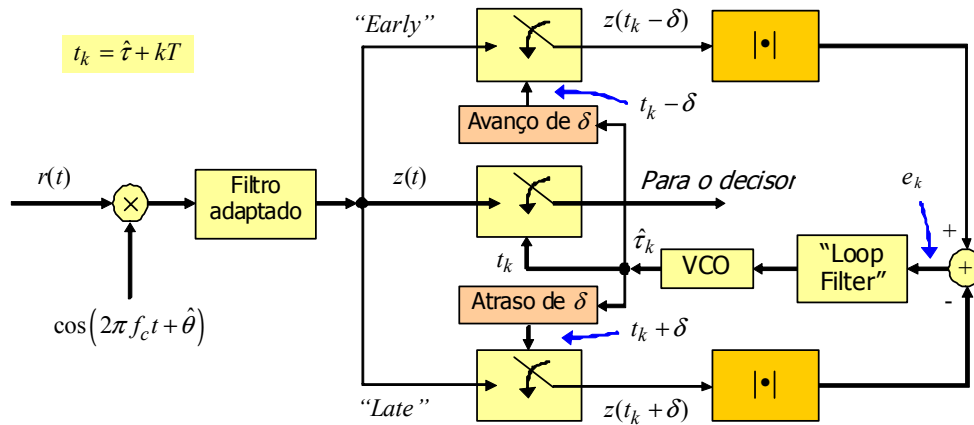


Fig. 35 O sincronizador “Early-Late” com filtro adaptado.

A Fig. 36 mostra três situações de sincronização. Quando o instante de amostragem central estimado está atrasado (isto é, quando ocorre depois do que devia, $\hat{\tau}_k > \tau$) o erro $e_k = |z(t_k - \delta)| - |z(t_k + \delta)|$ é positivo e o VCO do PLL aumenta a sua frequência, diminuindo assim $\hat{\tau}_k$. Se, pelo contrário, o relógio estiver adiantado ($\hat{\tau}_k < \tau$) o erro é negativo, o VCO diminui a sua frequência e $\hat{\tau}_k$ é aumentado.

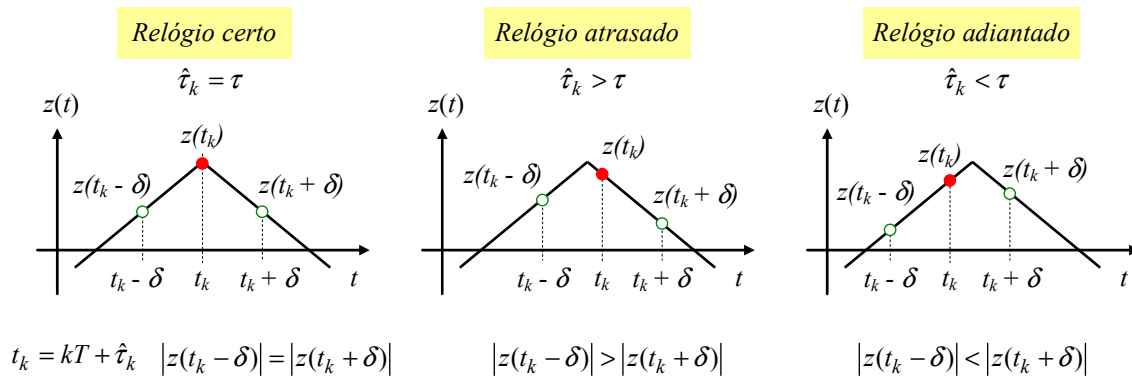


Fig. 36 O método “Early-Late” com filtro adaptado.

5.3 O método de sincronização de símbolo de Gardner

Os métodos de sincronização de símbolo em tempo discreto prestam-se naturalmente à sua aplicação em processadores digitais. Um dos mais usados foi proposto por Floyd Gardner em 1986 [10]. A Fig. 37 apresenta o respectivo diagrama de blocos para sinais bidimensionais (com PSK binário basta um dos ramos). Os amostradores da figura funcionam a uma cadência fixa de relógio superior a $1/T$ independente da temporização de símbolos e entregam as amostras aos interpoladores. Cada um destes produz duas amostras por símbolo, de $T/2$ em $T/2$ segundos, interpoladas nos instantes ditados pelo controlador, incluindo uma amostra no instante intermédio, $t_{k-1/2} = \frac{t_{k-1} + t_k}{2}$, entre cada dois instantes de símbolo consecutivos, t_k e t_{k-1} . Estes estão separados de cerca de T segundos. As amostras interpoladas nos referidos instantes t_{k-1} , $t_{k-1/2}$ e t_k , que não coincidem necessariamente com as amostras de saída dos amostradores, são aplicadas a um *Detector de Erro de Temporização*⁹ o qual, de acordo com o algoritmo apresentado a seguir, ajusta iterativamente a temporização dos símbolos. Os decimadores por 2 servem simplesmente para repor a sequência de amostras à taxa de símbolos, desprezando as amostras interpoladas nos instantes $t_{k-1/2}$.

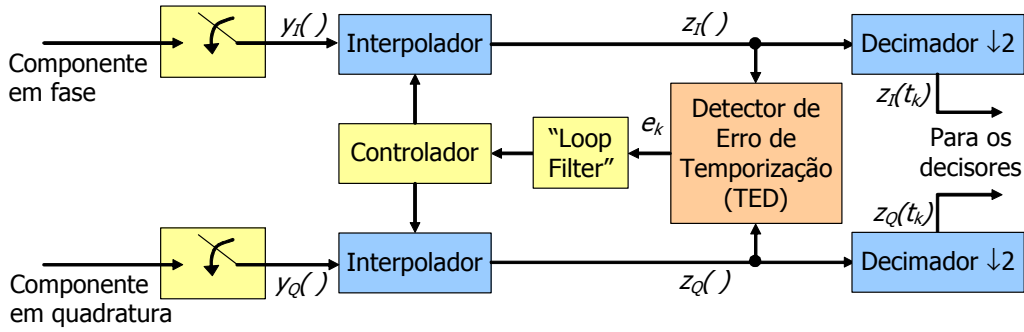


Fig. 37 Diagrama de blocos do método de sincronização de símbolo de Gardner.

O algoritmo de sincronização de Gardner é um método simples, com realimentação, independente da recuperação da fase da portadora. Vejamos como funciona em BPSK. Nesse caso o diagrama de blocos fica reduzido ao ramo em fase e as amostras z_I podem passar a ser designadas por z , para simplificar a notação. Durante cada símbolo o interpolador produz duas amostras: uma no próprio instante de símbolo t_k e outra aproximadamente $T/2$ segundos antes, no instante $t_{k-1/2}$. Em concreto, os símbolos de ordem $k-1$ e k são obtidos nos instantes

$$t_{k-1} = (k-1)T + \hat{\tau}_{k-1}$$

$$t_k = kT + \hat{\tau}_k,$$

respectivamente, e a amostra intermédia é interpolada no instante

$$t_{k-1/2} = \frac{t_{k-1} + t_k}{2} = kT - \frac{T}{2} + \frac{\hat{\tau}_{k-1} + \hat{\tau}_k}{2}.$$

Com base nos valores das amostras nesses três instantes o Detector de Erro de Temporização gera o erro

$$e_k = [z(t_k) - z(t_{k-1})]z(t_{k-1/2}),$$

usado depois para actualizar o próximo instante de interpolação de acordo com o algoritmo

$$t_{k+1} = t_k + T - \mu T e_k,$$

onde $\mu > 0$ é o passo de adaptação. Por conseguinte, a estimativa do atraso é actualizada segundo

$$\hat{\tau}_{k+1} = \hat{\tau}_k - \mu T e_k.$$

Este método de sincronização tem memória, portanto, pois o erro e_k no instante t_k depende do valor da amostra no instante t_{k-1} .

A Fig. 38, acompanhada da Fig. 37 com um só ramo, ajudará a compreender o método em BPSK. A sinalização é mais uma vez antipodal e rectangular e a saída do filtro adaptado, com a forma triangular habitual, é amostrada com uma frequência de amostragem muito maior que $1/T$, obtendo-se a onda $y(t)$ – a qual é, de facto, uma onda amostrada embora na figura esteja representada por traços contínuos¹⁰. Essa onda amostrada é aplicada ao interpolador que, por sua vez, produz as novas amostras $z(\cdot)$. Na situação à esquerda a temporização de símbolo está correcta e, portanto, a amostra intermédia é nula; logo, a saída do detector TED também é nula. Na situação central o “relógio” do interpolador está atrasado e $e_k > 0$; sendo assim, o próximo instante de interpolação, t_{k+1} , deve ser antecipado de modo a ocorrer menos de T segundos após t_k ; por fim, na situação à direita a interpolação $z(t_k)$ ocorreu cedo demais fazendo com que o erro e_k seja negativo e que o próximo instante de interpolação deva ser retardado.

¹⁰ O amostrador pode ser dispensado se a saída do filtro adaptado já for amostrada por natureza.

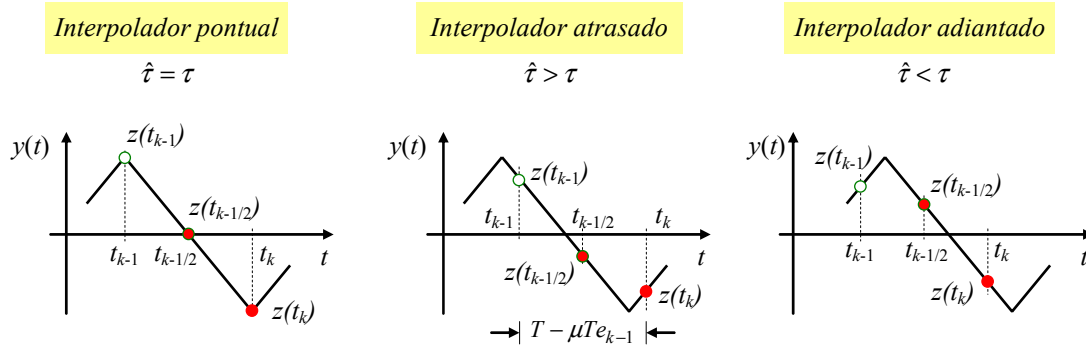


Fig. 38 Amostras interpoladas envolvidas no método de sincronização de Gardner.

O algoritmo funciona da mesma forma com sinais bidimensionais, como QPSK, exceptuando o bloco TED que, tendo em conta os dois ramos da Fig. 37, produz o erro

$$e_k = [z_I(t_k) - z_I(t_{k-1})]z_I(t_{k-1/2}) + [z_Q(t_k) - z_Q(t_{k-1})]z_Q(t_{k-1/2}).$$

Verificámos, portanto, que, tal como sucedia com o sincronizador “Early-Late”, ocorrem as seguintes situações relativamente a e_k :

$$e_k = \begin{cases} < 0 & \text{cedo demais} & \hat{\tau} \text{ deve aumentar} \\ = 0 & \text{pontual} & \hat{\tau} \text{ deve manter-se constante} \\ > 0 & \text{tarde demais} & \hat{\tau} \text{ deve diminuir} \end{cases}$$

Refira-se, para terminar, que o método de Gardner apresenta várias vantagens, entre as quais se salientam as duas seguintes:

- Só precisa de duas amostras por símbolo (uma das quais serve para a própria detecção dos dados).
- É independente da recuperação da portadora (pois o erro não depende do rigor da estimação de θ).

Exemplo 6:

Consideremos a Fig. 39. A saída do Detector de Erro de Estimação de Gardner vale

$$\begin{aligned} e_k &= [z(t_k) - z(t_{k-1})]z(t_{k-1/2}) = \\ &= [0,8 - (-0,7)] \times (-0,3) = \\ &= -0,45 \end{aligned}$$

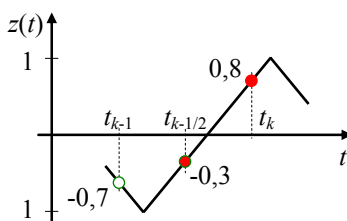


Fig. 39 Amostras obtidas no sincronizador de Gardner.

Sendo e_k negativo isso significa que o instante de interpolação t_k ocorreu cedo demais, ou seja, o “relógio” do interpolador está adiantado; logo, o próximo instante de interpolação deverá acontecer depois de $t = t_k + T$.

6. Conclusão

Foram abordadas as principais técnicas de sincronização usadas em modulações digitais. O tema é negligenciado em muitos livros de texto, mesmo nos mais referidos e reverenciados, que a ele dedicam meia dúzia de páginas, se tanto. Na verdade, ao serem apresentadas técnicas de detecção e desmodulação digitais é normalmente pressuposto que o sincronismo de símbolo e a recuperação de fase da portadora foram adquiridos adequadamente. Contudo, se o não tiverem sido, os resultados podem ser desastrosos, como vimos. A sincronização do receptor é, assim, uma tarefa crítica importante.

Neste texto considerámos que a recuperação da fase da portadora e a temporização de símbolo são realizadas separadamente. Encontrámos métodos de recuperação de fase que envolvem realimentação e têm na sua base o circuito PLL, como o tão usado “Costas Loop”, e vimos como proceder à estimação de máxima verosimilhança da fase em tempo contínuo e em tempo discreto, e neste caso com e sem a ajuda de dados ou decisões; tratámos também de métodos de temporização de símbolos em tempo contínuo, com dispositivo quadrático, e métodos em tempo discreto com os sincronizadores de Gardner e “Early-Late”. Enfim, observámos também a deterioração da probabilidade de erro na presença de erros de sincronismo.

Uma última palavra para referir uma relativamente nova tendência na sincronização: o aproveitamento do conceito, muito rico, de *processamento turbo*, surgido após a invenção dos códigos turbo em 1993 e vastamente aplicado em áreas francamente diversas que incluem, entre outras, precisamente a recuperação de temporização, mas agora realizada à luz da descodificação iterativa [11].

7. Bibliografia recomendada

- [1] Bougeard, S., Héliard, J.-F. e Siaud, I., “Performance Optimization of High Order QAM in Presence of Phase Noise and AWGN: Application to a Decision-Directed Frequency Synchronization System”, *Wireless Personal Communications*, Vol. 37, nº 1-2, Abril de 2006, pp. 123-138.
- [2] Meyr, H. e Ascheid, G., *Synchronization in Digital Communications*, Wiley, 1990.
- [3] Cioffi, John M., *EE 379A (Digital Communication: Signal Processing) Course Reader*, Stanford University, <http://www.stanford.edu/class/ee379a/>. Acedido em Fevereiro de 2007.
- [4] Haykin, Simon, *Communication Systems*, 4ª Ed., Wiley, 2001.
- [5] Mengali, Umberto e D’Andrea, Aldo N., *Synchronization Techniques for Digital Receivers*, Plenum Press, 1997.
- [6] Moeneclaey, Marc e de Jonghe, Geert, “ML-Oriented NDA Carrier Synchronization for General Rotationally Symmetric Signal Constellations”, *IEEE Transactions on Communications*, Vol. COM-42, nº 8, Agosto de 1994, pp. 2531-2533.
- [7] Viterbi, A. J. e Viterbi, A. M., “Nonlinear estimation of PSK-modulated carrier phase with application to burst digital transmission”, *IEEE Transactions on Information Theory*, Vol. 29, nº 4, Julho de 1983, pp. 543-551.
- [8] Proakis, John G., *Digital Communications*, 3ª Ed., McGraw-Hill, 1995.
- [9] Sklar, Bernard, *Digital Communications: Fundamentals and Applications*, 2ª Ed., Prentice-Hall, 2001.
- [10] Gardner, F. M., “A BPSK/QPSK Timing-Error Detector for Sampled Receivers”, *IEEE Transactions on Communications*, Vol. COM-34, nº 5, Maio de 1986, pp. 423-429.
- [11] Barry, J. R., Kavcic, A., McLaughlin, S. W., Nayak, A. e Zeng, W., “Iterative Timing Recovery”, *IEEE Signal Processing Magazine*, Vol. 21, nº 1, Janeiro de 2004, pp. 89-102.