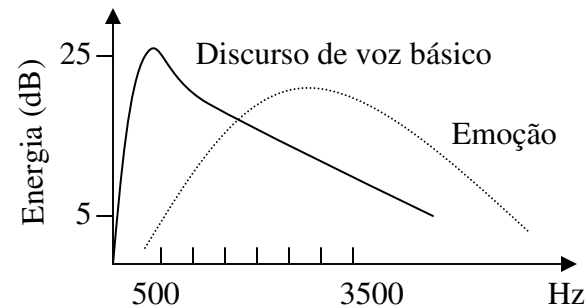


Características do sinal de voz

- Análise na frequência:
 - a voz apresenta um conteúdo espectral que vai de 20 Hz a 20 kHz;
 - os sons vozeados ou nasais (e.g. vogais e algumas consoantes j, l, m) apresentam um espectro discreto com uma frequência fundamental de 100 a 200 Hz nos homens e 200 a 400 Hz nas mulheres;
 - os sons não vozeados (e.g. f, s, p, ch) que são gerados pelo fluxo de ar na boca modulado pelos maxilares, língua e lábios apresentam uma variação aleatória. O seu espectro é contínuo;
 - as frequências mais baixas transportam a energia do sinal e as mais elevadas a emoção.



Características do sinal de voz (cont.)

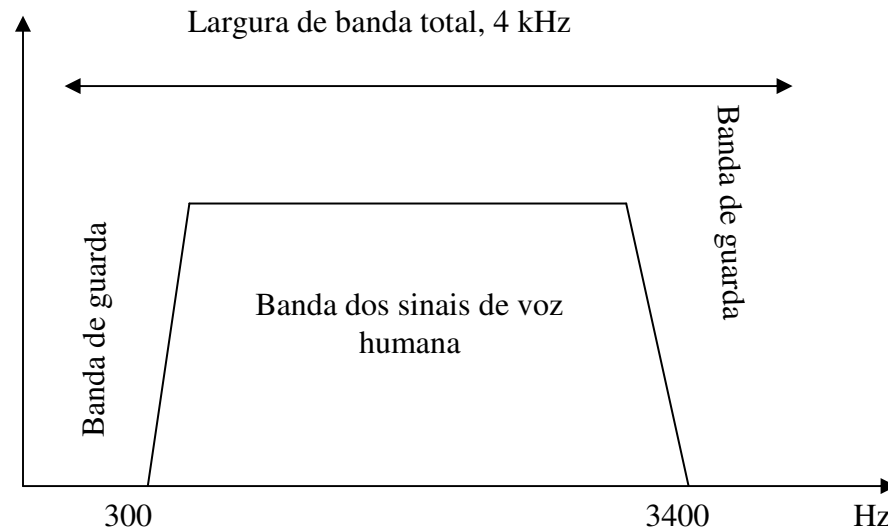
- Análise no tempo:
 - várias sílabas por segundo;
 - a fala concentra-se em intervalos de duração aleatória (com média de cerca de 1 seg.) separados por intervalos de duração aleatória (superior a 100 ms, quando se está a falar) \Rightarrow variação temporal bastante irregular e aleatória;
 - o sinal de voz só está presente, em média, em 40% do tempo. Pode-se aproveitar este facto para intercalar outras conversações (sistema TASI - time assignment speech interpolation).

Caracterização do sistema auditivo

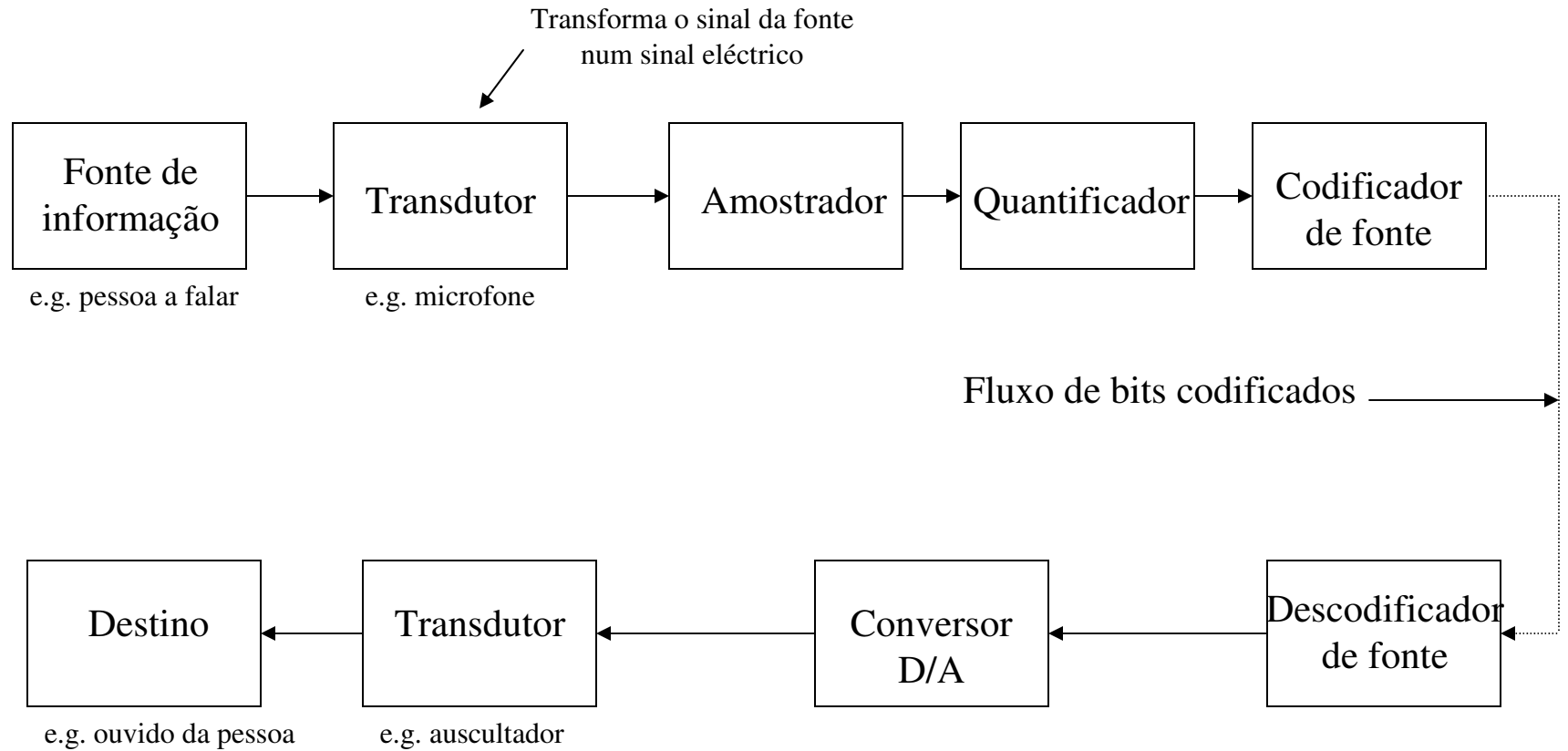
- Um indivíduo normal com idade compreendida entre os 18 e 25 anos é capaz de detectar sons puros entre 20 Hz e 20 kHz;
- Com a idade, o limite superior da frequência audível reduz-se significativamente, e.g. em média um homem de 65 anos tem a 8 kHz uma perda de sensibilidade de 40 dB;
- A sensibilidade do ouvido varia com a frequência e com a intensidade sonora. Este aspecto terá de se reflectir na análise do desempenho das redes telefónicas, em particular na medida da potência do ruído:
potência psofométrica (dBmp)
Ruído branco (0-4000Hz) -3.6dB; (300-3400Hz) -2.5dB
- O ouvido tem uma elevada gama dinâmica, com valores que podem ir acima dos 100 dB. Para uma boa reprodução basta valores da ordem dos 30 dB.

Banda de frequências normalizada para a voz

- As recomendações G.132 e G.151 do ITU-T indicam a banda atribuída ao sinal de voz de 300 - 3400 Hz;
- Nos EUA a banda de frequências atribuída para um canal de voz é 200 - 3200 Hz;
- Estas larguras de banda resultam de um compromisso entre o que os assinantes telefónicos pretendem e o que lhes pode ser fornecido economicamente.



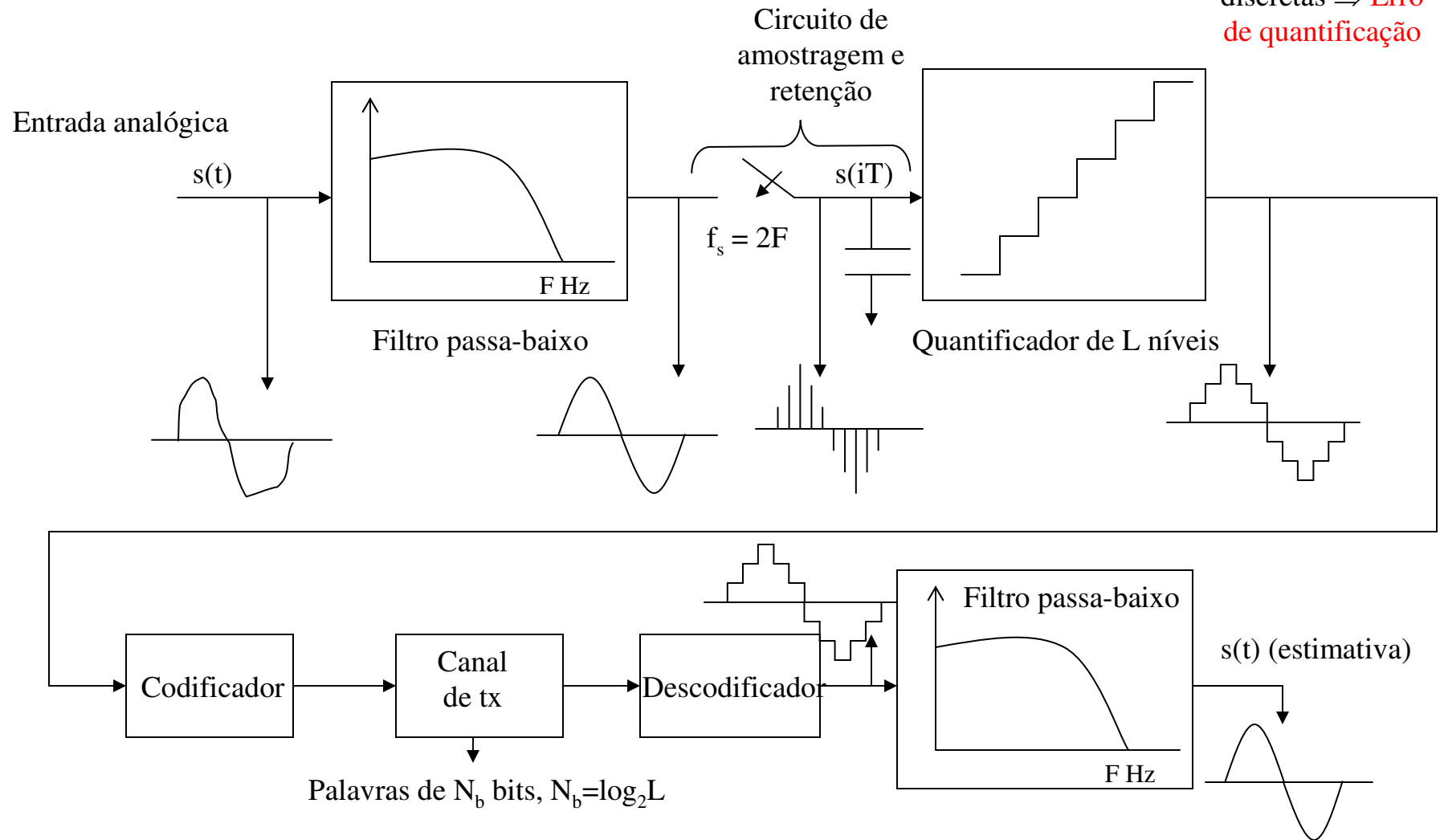
Transmissão digital de sinais analógicos



PCM - Pulse Code Modulation

- Diagrama de blocos -

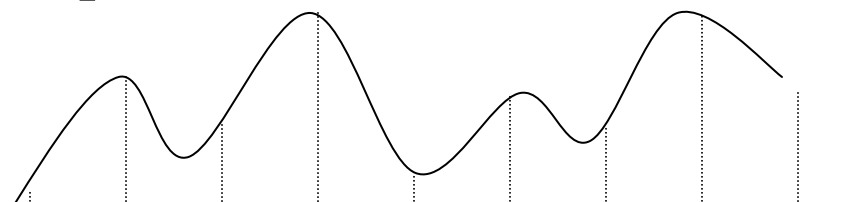
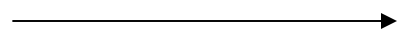
Conversão das amostras num conjunto de amplitudes discretas \Rightarrow Erro de quantificação



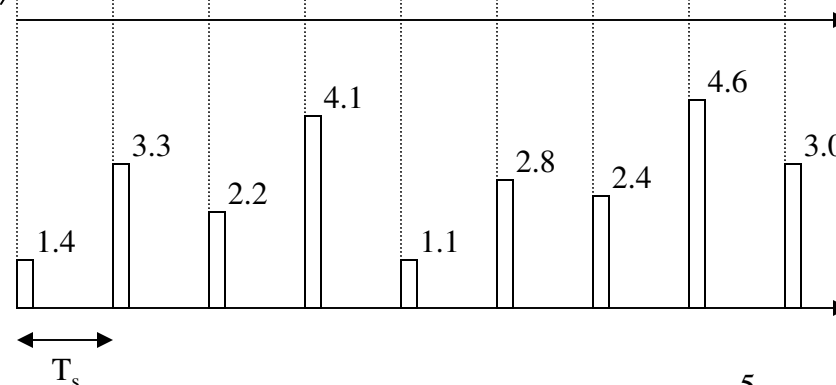
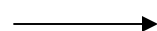
PCM - Pulse Code Modulation

- Exemplo -

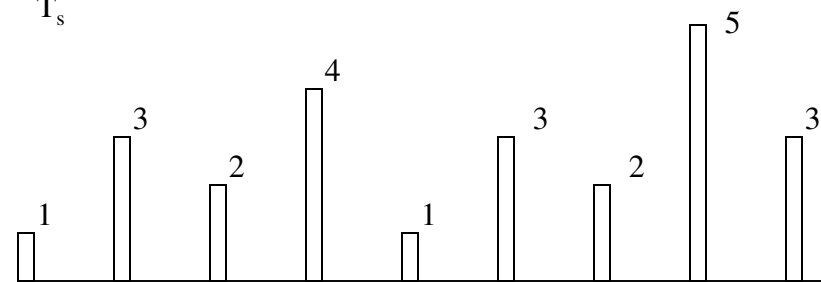
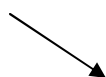
Sinal original



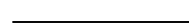
Resultado da amostragem



Resultado da quantificação pelo inteiro mais próximo

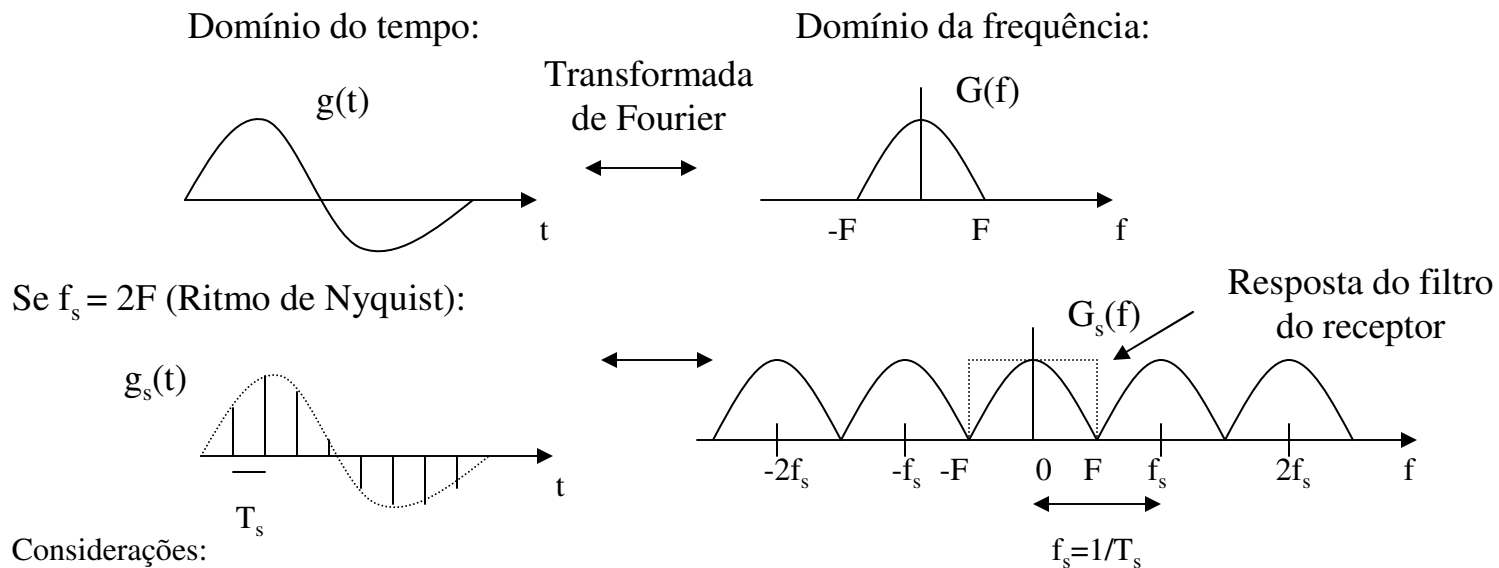


Resultado da codificação (palavra binária), i.e. sinal PCM



Amostragem

- A base do PCM começa com o Teorema da Amostragem:
 - Um sinal de banda limitada pode ser representado pelas suas amostras obtidas a um ritmo f_s que deve ser pelo menos igual ao dobro da frequência máxima presente no sinal (F), i.e. $f_s \geq 2F$.



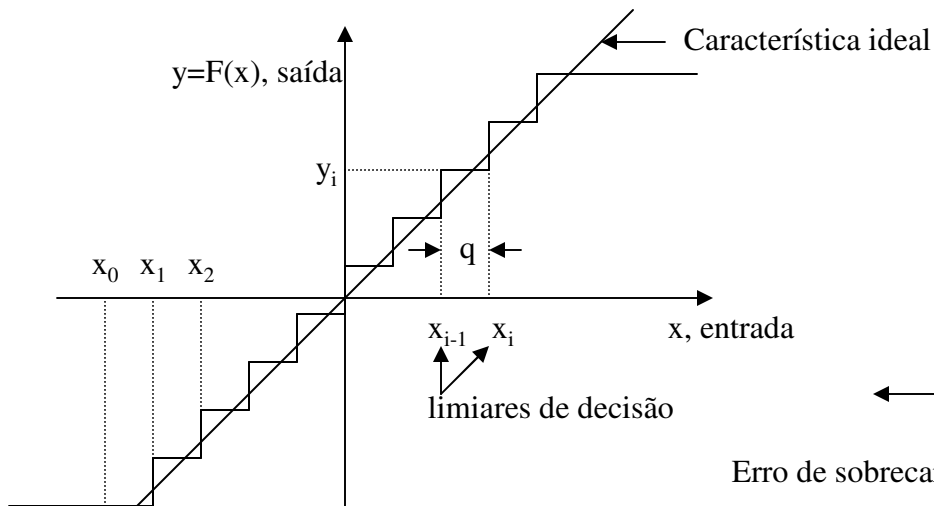
Considerações:

- Sinal $g(t)$ tem uma largura de banda finita (F Hz) - Na prática não se verifica \Rightarrow Aliasing
- As amostras são retiradas com pulsos de largura infinitesimal
- Filtro passa-baixo ideal

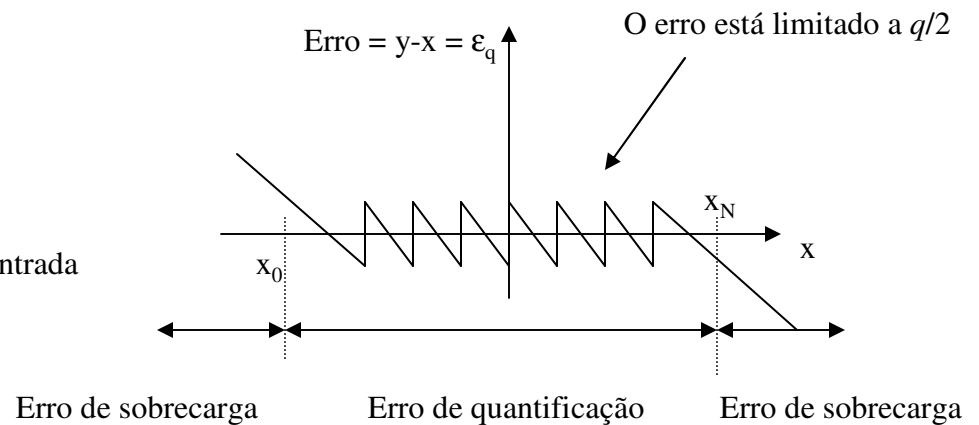
Quantificação

- A quantificação converte um sinal contínuo em amplitude num sinal discreto em amplitude. Notar que o processo de amostragem converte um sinal contínuo no tempo num sinal discreto no tempo - Pulse Amplitude Modulation (PAM).

Característica do quantificador linear ou uniforme:



Característica do erro:



NOTA: Quando o valor de entrada está entre x_{i-1} e x_i o quantificador irá produzir o valor y_i

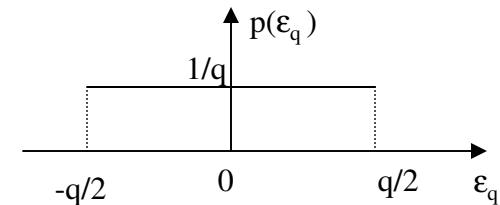
Ruído de quantificação

- O desempenho de um quantificador pode ser descrito pela relação sinal-ruído de quantificação;
- A potência de ruído de quantificação é descrita em termos estatísticos através do erro quadrático médio:

$$\langle \varepsilon_q^2 \rangle = \int_{-\infty}^{+\infty} \varepsilon_q^2 p(\varepsilon_q) d\varepsilon_q$$

- Para uma distribuição uniforme do erro em cada intervalo de quantificação de largura q ;

$$n_q = \langle \varepsilon_q^2 \rangle = \int_{-\infty}^{+\infty} \varepsilon_q^2 \frac{1}{q} d\varepsilon_q = \frac{q^2}{12}$$



- Para um quantificador linear, todos os intervalos de quantificação apresentam o mesmo erro quadrático médio.

Cálculo da relação sinal-ruído de quantificação

- Assume-se que o sinal de entrada é uma sinusóide com amplitude A , logo a potência média desse sinal vem $s = \langle x(t)^2 \rangle = A^2/2$

- A relação sinal-ruído de quantificação, em dB:

$$\frac{S}{N_q} = 10 \log_{10} \left(\frac{s}{n_q} \right) = 10 \log_{10} \left(\frac{A^2/2}{q^2/12} \right) = 7.78 + 20 \log_{10} \left(\frac{A}{q} \right)$$

- Para PCM uniforme não inferior a 26dB
- Número de intervalos de quantificação para uma gama de quantificação de $-A_{max}$ a A_{max} :

$$L = \frac{A_{max} - (-A_{max})}{q} = \frac{2A_{max}}{q}$$

- Número de bits por amostra: $L = 2^{N_b} \Leftrightarrow N_b = \log_2 L$

Cálculo da relação sinal-ruído de quantificação (cont.)

- Relação sinal-ruído de quantificação em termos do número de bits, N_b

$$\frac{S}{N_q} = 1.76 + 6.02N_b + 20\log_{10}\left(\frac{A}{A_{\max}}\right)$$

- Para um determinado n° de bits por amostra constante a relação sinal-ruído de quantificação depende da amplitude A do sinal a quantificar:
 - sinais com baixa amplitude têm uma relação S/N_q baixa, enquanto os sinais com amplitude elevada apresentam S/N_q elevadas;
 - sinais com elevadas amplitudes têm pouca probabilidade de ocorrer e os sinais com baixas amplitudes ocorrem mais frequentemente.
- ⇒ PCM uniforme é pouco eficiente.

Gama dinâmica

- Conceito: Relação entre a amplitude máxima, A_{max} , e a amplitude mínima, A_{min} , em que o sistema deve ser capaz de funcionar com a qualidade mínima

$$GD = 20 \log_{10} \left(\frac{A_{max}}{A_{min}} \right)$$

- Se se quiser assegurar uma qualidade S/N_q para toda a gama dinâmica, então o n° de bits necessário deve verificar

$$\frac{S}{N_q} = 1.76 + 6.02N_b - GD$$

- O sistema telefónico deve ser capaz de transmitir uma elevada gama de amplitudes, i.e. deve ter uma gama dinâmica elevada (30 dB é um valor típico).

Quantificação não-uniforme

- Com quantificação uniforme $\frac{S}{N_q} = 1.76 + 6.02N_b - GD$
 - gamas dinâmicas elevadas exigem um n° de bits por amostra, N_b , elevado para garantir uma S/N_q especificada;

Exemplo: $GD = 50$ dB, $S/N_q = 30$ dB \Rightarrow 13 bits/amostra

- a S/N_q resultante é demasiado elevada para sinais fortes;
- clientes diferentes são servidos com qualidade diferente.

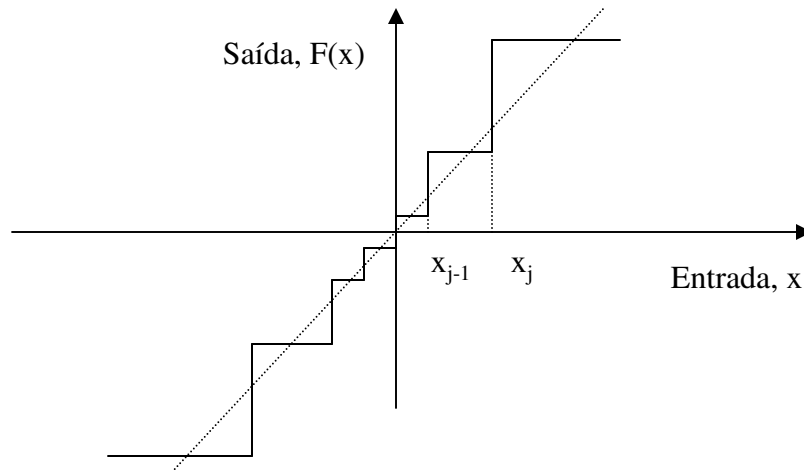
Solução: Quantificação não-uniforme

- Para se obter S/N_q independente da amplitude do sinal, o intervalo de quantificação deve ser proporcional à amplitude do sinal.

Quantificação não-uniforme (cont.)

- Solução: dividir a amplitude do sinal de entrada em intervalos não-uniformes, i.e. intervalos de quantificação mais largos para os sinais de amplitudes elevadas e intervalos mais estreitos para amplitudes baixas
 $\Rightarrow S/N_q$ constante para uma característica de quantificação apropriada.

Característica do quantificador não-uniforme:



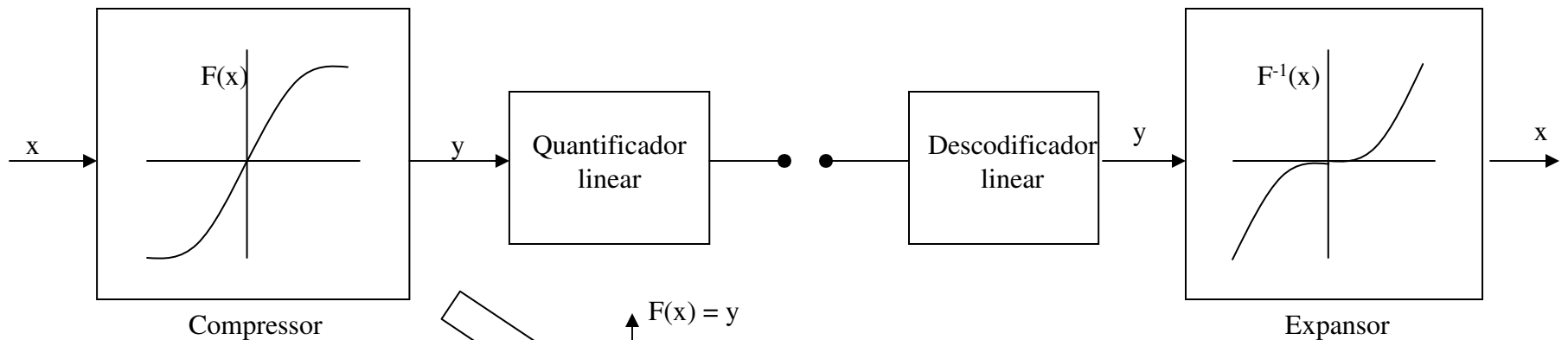
Como se realiza esta função ?

- Possível solução: compressão das amostras seguida de quantificação linear.

Implementação do quantificador não-uniforme

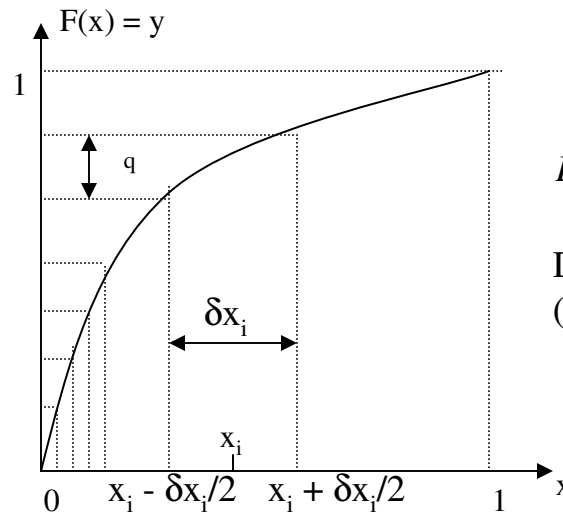
Emissor:

Receptor:



Característica normalizada do compressor (só valores positivos):

NOTA: $x_i - 1/2\delta x_i \leq x \leq x_i + 1/2\delta x_i$
 Este sinal x vai ser representado pela amplitude quantificada x_i



L níveis na gama de -1 a 1: $q = 2/L$

Declive da característica do compressor (L elevado, $q \rightarrow 0$, $\delta x_i \rightarrow 0$)

$$\delta x_i = \frac{2}{L} \left(\frac{dx}{dy} \right)$$

Potência do ruído de quantificação não-uniforme

- O erro quadrático médio total é dado, em termos estatísticos, por

$$n_q = \langle \varepsilon_q^2 \rangle = \sum_{i=1}^L \underbrace{\int_{x_i - \delta x_i/2}^{x_i + \delta x_i/2} (x_i - x)^2 p(x) dx}_{\text{Contribuição de cada intervalo de quantificação}} \approx \sum_{i=1}^L p(x_i) \int_{-\delta x_i/2}^{\delta x_i/2} \varepsilon_i^2 d\varepsilon_i \longrightarrow \varepsilon_i \text{ é o erro de quantificação do intervalo } i$$

Assume-se que a fdp de x é constante em cada intervalo

$$n_q = \sum_{i=1}^L p(x_i) \frac{(\delta x_i)^3}{12} = \sum_{i=1}^L \frac{(\delta x_i)^2}{12} \underbrace{p(x_i) \delta x_i}_{\text{Probabilidade do sinal } x \text{ estar no } i\text{-ésimo intervalo}} = \frac{1}{3L^2} \sum_{i=1}^L \left(\frac{dx}{dy} \right)^2 p(x_i) \delta x_i$$

$$\delta x_i = \frac{2}{L} \left(\frac{dx}{dy} \right)$$

NOTA: no caso da quantificação uniforme $\delta x_i = q$ resultando em $n_q = q^2/12$

O ruído de quantificação depende da estatística do sinal analógico a discretizar

No caso em que L é elevado ($\delta x_i \rightarrow 0$):

$$n_q = \frac{1}{3L^2} \int_{-1}^1 \left(\frac{dx}{dy} \right)^2 p(x) dx$$

Relação sinal-ruído de quantificação

- Potência do sinal (também depende da estatística do sinal):

$$s = \langle x^2 \rangle = \int_{-1}^1 x^2 p(x) dx$$

- Relação sinal-ruído de quantificação:

$$\frac{s}{n_q} = 3L^2 \frac{\int_{-1}^1 x^2 p(x) dx}{\int_{-1}^1 \left(\frac{dx}{dy}\right)^2 p(x) dx}$$

- Tem interesse em definir-se uma relação sinal-ruído de quantificação independente (da estatística) do sinal de entrada:

$$\frac{dx}{dy} = kx \Rightarrow \frac{s}{n_q} = \frac{3L^2}{k^2}$$

Relação sinal-ruído proporcional ao quadrado do nº de níveis \Rightarrow melhoria de 6 dB por cada bit a mais na codificação

Resolução da eq. diferencial:

$$\frac{dy}{dx} = \frac{1}{kx} \quad \boxed{x=1 \Rightarrow y=1} \quad y = 1 + \frac{1}{k} \ln x \rightarrow$$

Compressão logarítmica: característica irrealizável devido à assíntota vertical para $x = 0$

Ganho de compressão / Vantagem de compressão

- Ganho de compressão, relativamente à quantificação uniforme, g_c
 - Relação sinal-ruído de quantificação não-uniforme:

$$\left(\frac{s}{n_q}\right)_{n-u} = 3L^2 s / \int_{-1}^1 \left(\frac{dx}{dy}\right)^2 p(x) dx$$

- Relação sinal-ruído de quantificação uniforme, $dy/dx = 1$: $\left(\frac{s}{n_q}\right)_u = 3L^2 s$

$$\Rightarrow \boxed{g_c = \frac{(s/n_q)_{n-u}}{(s/n_q)_u} = 1 / \int_{-1}^1 \left(\frac{dx}{dy}\right)^2 p(x) dx} \rightarrow \text{Depende (da estatística) do sinal e da característica do compressor}$$

- Vantagem de compressão, $v_c =$ ganho de compressão para sinais (muito) fracos

$$\lim_{x \rightarrow 0} p(x) = \delta(x) \Rightarrow v_c = \lim_{x \rightarrow 0} g_c = \lim_{x \rightarrow 0} \left(\frac{dy}{dx}\right)^2 \rightarrow \text{Depende unicamente da característica do compressor}$$

PCM não-linear (leis A e μ)

- Duas implementações de características de compressão logarítmicas
 - na Europa: lei A

$$y = \begin{cases} \operatorname{sgn}(x) \left[\frac{A|x|}{1 + \ln(A)} \right] & 0 \leq |x| \leq \frac{1}{A} \\ \operatorname{sgn}(x) \left[\frac{1 + \ln(A|x|)}{1 + \ln(A)} \right] & \frac{1}{A} \leq |x| \leq 1 \end{cases} \Rightarrow \begin{cases} V_c = 10 \log_{10} \left(\frac{A}{1 + \ln(A)} \right)^2 \\ V_c = 24 \text{ dB } (A = 87.6) \end{cases}$$

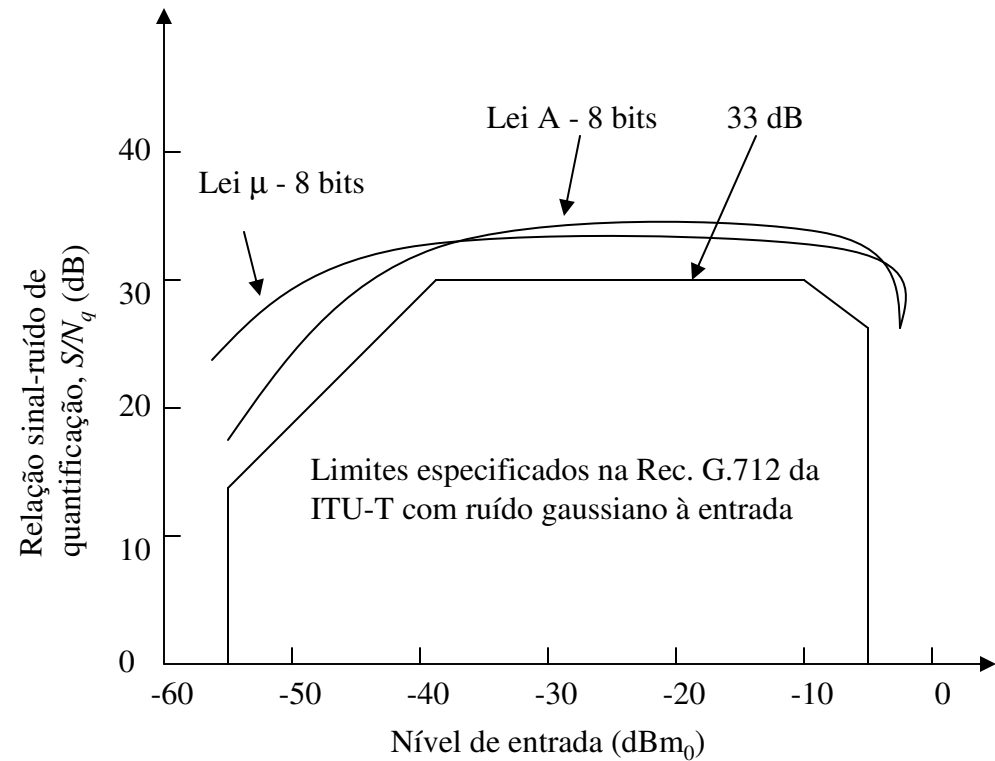
Melhoria de 24 dB na zona das baixas amplitudes em relação à quantificação uniforme

- nos EUA e Japão: lei μ (μ define o grau de compressão e $\mu = 255$ é um valor típico)

$$y = \operatorname{sgn}(x) \frac{\ln(1 + \mu|x|)}{\ln(1 + \mu)} \Rightarrow \begin{cases} V_c = 10 \log_{10} \left(\frac{\mu}{\ln(1 + \mu)} \right)^2 \\ V_c = 33.3 \text{ dB } (\mu = 255) \end{cases}$$

Quantificação não uniforme usada para a voz

- Numa GD de 40 dB a lei μ tem uma S/N_q mais uniforme que a lei A.



Técnicas de quantificação não-uniforme

- Compressão e expansão logarítmica analógica (e.g. díodos) seguida de quantificação uniforme com 8 bits,
 - dificuldade em garantir a reciprocidade exacta das características de compressão e expansão.
- Compressão aproximada por uma característica segmentada (e perfeitamente adaptada ao processamento digital),
 - PCM segmentado de 13 segmentos usando a lei A;
 - PCM segmentado de 15 segmentos usando a lei μ .

PCM segmentado de 13 segmentos (lei A, $A = 87.6$, 8 bits)

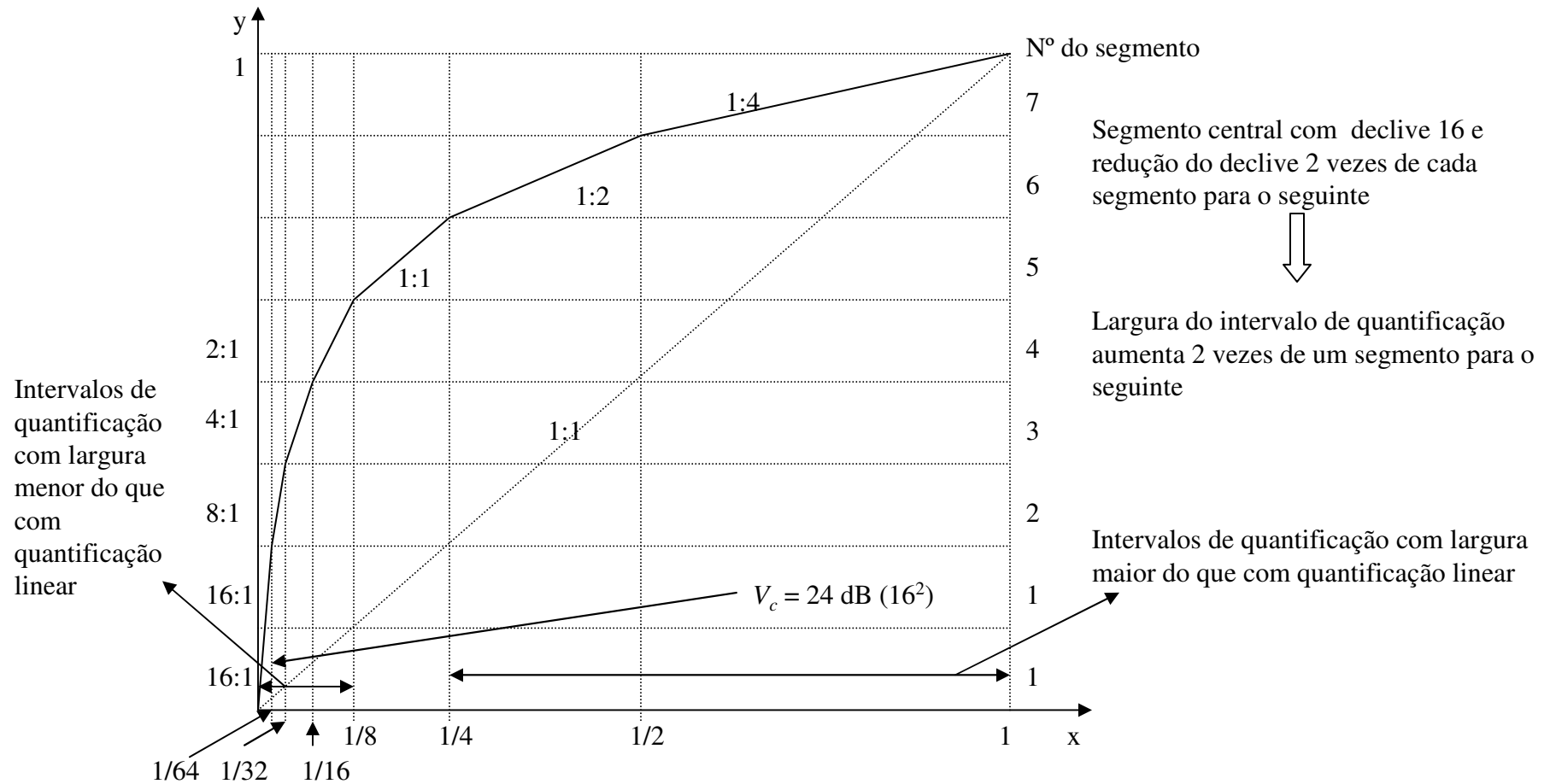
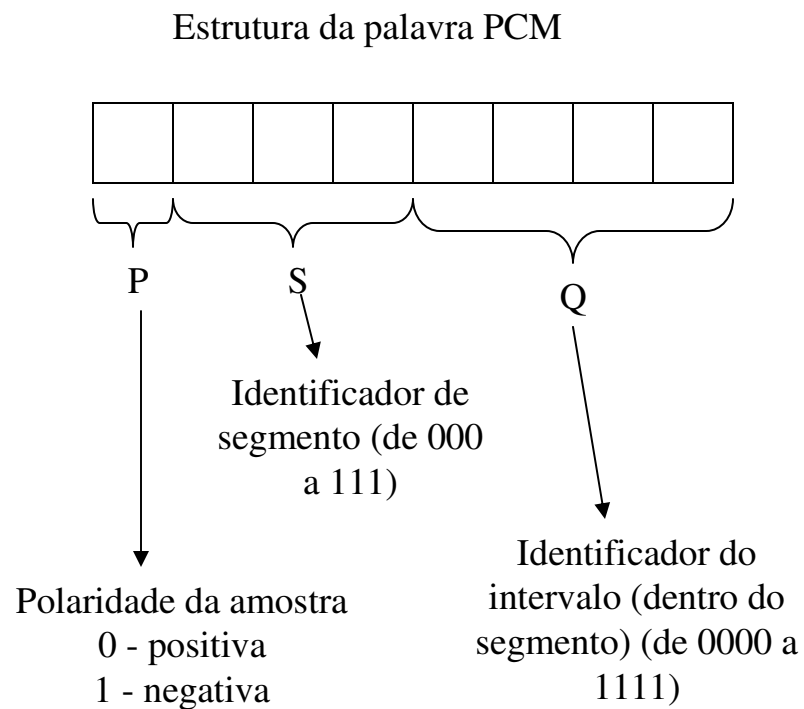


Tabela de codificação da lei A segmentada

Nº do segmento	gama do sinal	dimensão do passo	código do segmento	código de quantificação
1	0-2	2	000	0000
	2-4			0001
2	30-32	4	010	1111
	32-34			0000
3	62-64	8	011	1111
	64-68			0000
4	124-128	16	100	1111
	128-136			0000
5	248-256	32	101	1111
	256-272			0000
6	496-512	64	110	1111
	512-544			0000
7	992-1024	128	111	1111
	1024-1088			0000
	1984-2048			1111
	2048-2176			0000
	3968-4096			1111



Requisitos do sinal de voz / Parâmetros típicos do PCM para a voz

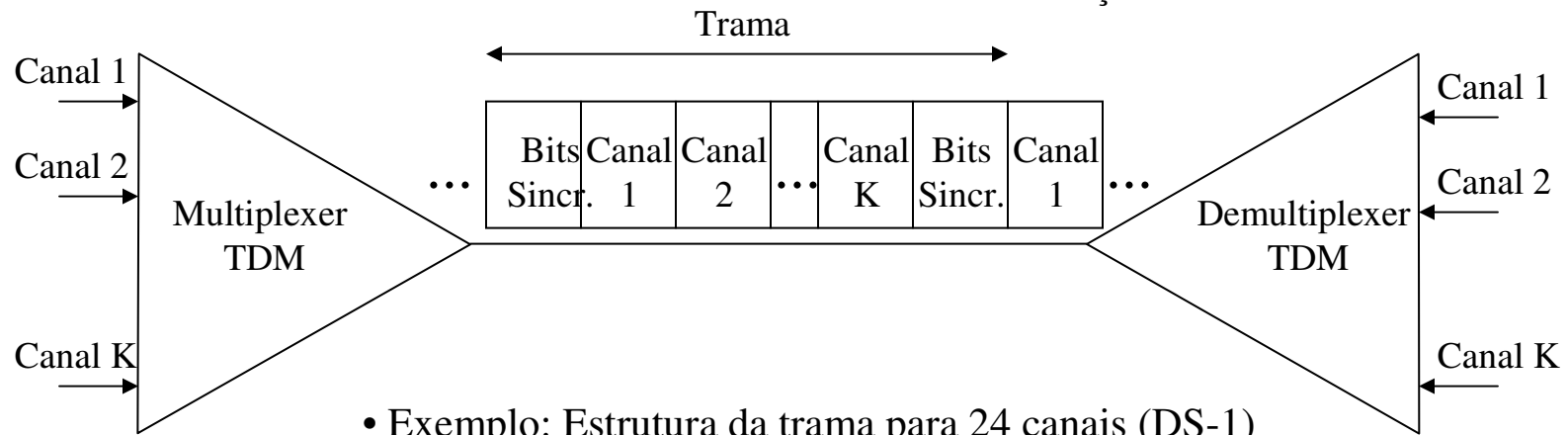
- Requisitos para transmissão de voz:
 - Testes demonstraram que para garantir uma boa qualidade de transmissão de voz é necessário garantir

$$\left(\frac{S}{N_q} \right)_{\text{dB}} \geq 35 \text{ dB}$$

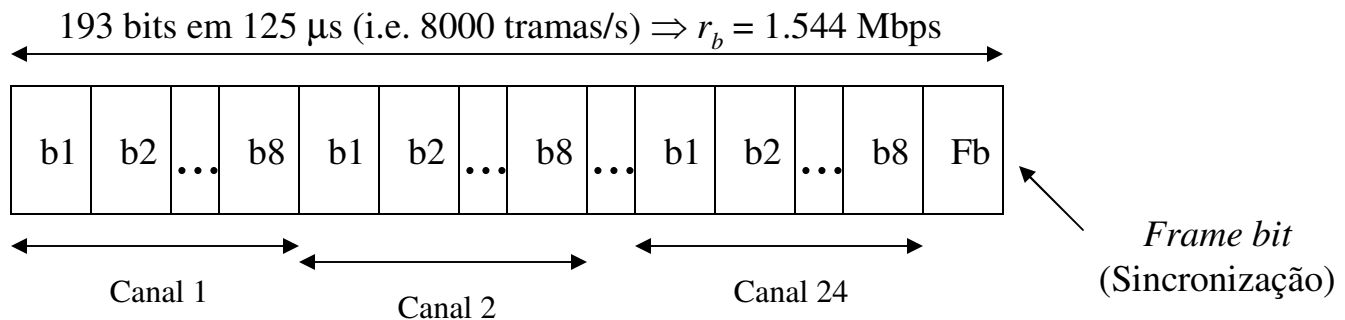
- Esta condição deve ser respeitada para uma gama dinâmica da ordem dos 30 dB.
- Características do sistema PCM para a voz:
 - Frequência de amostragem: 8000 amostras/s
 - Quantificação não uniforme com $L = 256$ níveis;
 - Compressão segundo a lei A com 13 segmentos (Europa) ou lei μ (EUA e Japão) com 15 segmentos;
 - palavras PCM de 8 bits;
 - Ritmo binário: 64 kbit/s.

Multiplexagem no tempo (Time-Division Multiplexing, TDM)

é o mais pequeno grupo de bits contendo pelo menos 1 amostra de cada canal mais os bits de sincronização



• Exemplo: Estrutura da trama para 24 canais (DS-1)



Ver Carlson,
Cap. 12.5 !

Multiplexagem de canais PCM e hierarquias adoptadas pelo ITU-T

