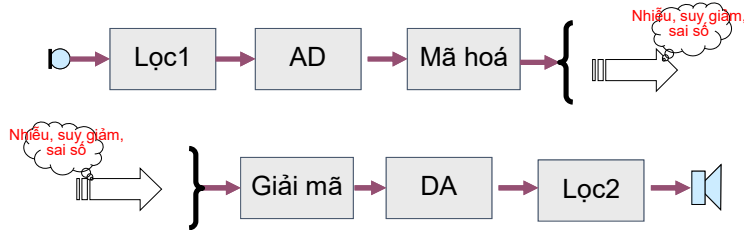


2. Mã hóa tiếng nói

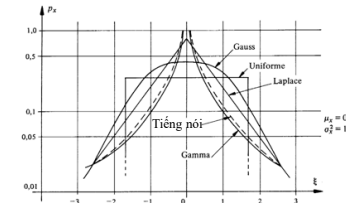
- Dãy thao tác mã hoá và giải mã



Một số tính chất thống kê của tín hiệu tiếng nói

- Mật độ xác suất

N_ξ : số lượng mẫu $x(n)$
có biên độ trong
khoảng $[\xi - \Delta\xi/2, \xi + \Delta\xi/2]$
 $n \in [-N, \dots, N]$
 x ergodic và dừng



$$p_x(\xi) = \lim_{\substack{N \rightarrow \infty \\ \Delta\xi \rightarrow 0}} [N_\xi / (2N + 1)]$$

Giá trị trung bình và phương sai

- Giá trị trung bình của tín hiệu dừng

$$\mu_x = \int_{-\infty}^{\infty} \xi p_x(\xi) d\xi = \lim_{N \rightarrow \infty} \frac{1}{2N + 1} \sum_{n=-N}^N x(n)$$

với tín hiệu tiếng nói $\mu_x = 0$

- Phương sai

$$\sigma_x^2 = \int_{-\infty}^{\infty} \xi^2 p_x(\xi) d\xi = \lim_{N \rightarrow \infty} \frac{1}{2N + 1} \sum_{n=-N}^N x^2(n)$$

Lượng tử tức thời (không nhớ)

- Luật lượng tử $y = Q(x)$ được định nghĩa:

- $(L + 1)$ mức tín hiệu $x(0), x(1), \dots, x(L)$
- L mức lượng tử hoá

- Mỗi mức lượng tử hoá biểu diễn bằng từ b bit

$$L = 2^b.$$

- Sai số lượng tử (tạp âm lượng tử) $e_q = Q(x) - x$

- Bước lượng tử : hiệu 2 mức tín hiệu kề nhau

$$\Delta(i) = x(i) - x(i - 1)$$

- Thông lượng $I = bFs$ (bit/s). Fs : tần số lấy mẫu

Thông lượng

- Tín hiệu lượng tử 8 bit (256 mức), $F_s = 8 \text{ kHz} \rightarrow$ Thông lượng = 64 kbit/s
- Tín hiệu lượng tử 16 bit (65536 mức),
 $F_s = 16 \text{ kHz} \rightarrow$ Thông lượng = 256 kbit/s ,
1 giờ tiếng nói $\sim 100 \text{ Mbyte}$
- Cần phải mã hoá tín hiệu tiếng nói (MPEG, GSM, G723, ...) để truyền tiếng nói trên mạng hoặc lưu trữ



66

66

Thông lượng

Tần số lấy mẫu (kHz)	Số bit cho 1 mẫu	Thông lượng kbit/s	Dung lượng / phút (kbyte)	Lĩnh vực
48	16	768	11520	Ghi âm chuyên nghiệp
44,1	16	705,6	10584	CD Audio
32	16	512	7680	Radio FM
22	12	264	3960	Radio AM
8	8	64	960	Điện thoại

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

67

67

Lượng tử đều

- Tổng quát, bước lượng tử là hàm của biên độ tín hiệu x (lượng tử không đều) \rightarrow đơn giản nhất là lượng tử đều.
- Lượng tử đơn cực: Tín hiệu tương tự biến thiên từ 0 von đến một giá trị dương nào đó.
- Lượng tử lưỡng cực: Tín hiệu tương tự biến thiên từ giá trị âm đến giá trị dương nào đó.
- x_{max}, x_{min} : giá trị cực đại và cực tiểu của tín hiệu tương tự x

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

68

68

Lượng tử đều

- L : Số mức lượng tử, b : số bit cho một mức lượng tử dùng trong ADC. $L = 2^b$
- Bước lượng tử $\Delta = (x_{max} - x_{min}) / L$
- i : chỉ số tương ứng với mã nhị phân
 $i = \text{round}((x - x_{min}) / \Delta)$
- x_q : mức lượng tử
 $x_q = x_{min} + i\Delta, i = 0, 1, \dots, L - 1$
- e_q : sai số lượng tử $e_q = x_q - x$

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

69

69

Lượng tử đều

• Đơn cực

$$x_{\min} = 0$$

$$x_{\max} = 8\Delta$$

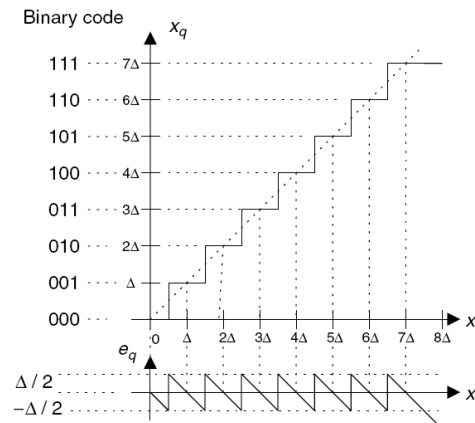
$$b = 3$$

$$L = 8$$

$$x_q = 0 + i\Delta,$$

$$i = 0, 1, \dots, L-1$$

$$-\Delta/2 \leq e_q \leq \Delta/2$$



Lượng tử đều

Bảng lượng tử của bộ lượng tử đơn cực 3 bit, $x_{\min} = 0$

x_{\max} = giá trị điện áp cực đại

Binary Code	Quantization Level x_q (V)	Input Signal Subrange (V)
0 0 0	0	$0 \leq x < 0.5\Delta$
0 0 1	Δ	$0.5\Delta \leq x < 1.5\Delta$
0 1 0	2Δ	$1.5\Delta \leq x < 2.5\Delta$
0 1 1	3Δ	$2.5\Delta \leq x < 3.5\Delta$
1 0 0	4Δ	$3.5\Delta \leq x < 4.5\Delta$
1 0 1	5Δ	$4.5\Delta \leq x < 5.5\Delta$
1 1 0	6Δ	$5.5\Delta \leq x < 6.5\Delta$
1 1 1	7Δ	$6.5\Delta \leq x < 7.5\Delta$

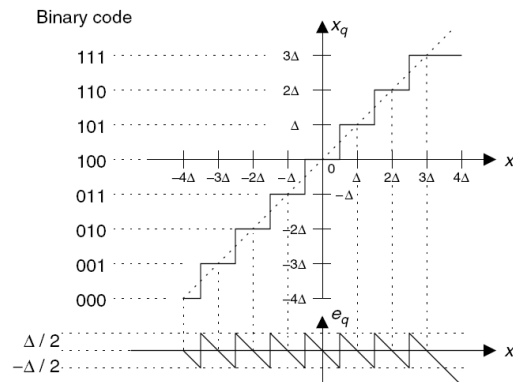
Lượng tử đều

• Lượng cực

$$x_{\min} = -4\Delta$$

$$x_{\max} = 4\Delta$$

$$b = 3$$



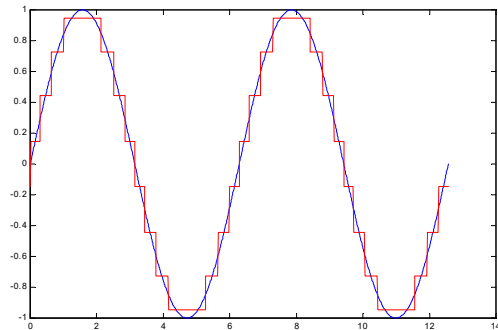
Lượng tử đều

• Bảng lượng tử của bộ lượng tử lưỡng cực 3 bit,

x_{\max} = giá trị điện áp cực đại, $x_{\min} = -x_{\max}$

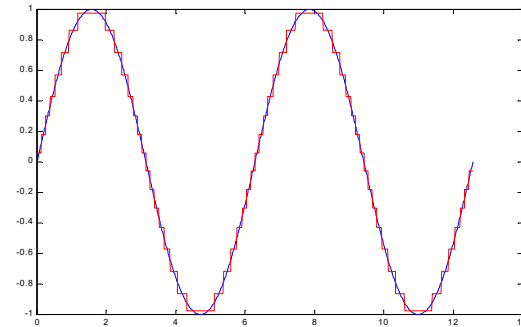
Binary Code	Quantization Level x_q (V)	Input Signal Subrange (V)
000	-4Δ	$-4\Delta \leq x < -3.5\Delta$
001	-3Δ	$-3.5\Delta \leq x < -2.5\Delta$
010	-2Δ	$-2.5\Delta \leq x < -1.5\Delta$
011	$-\Delta$	$-1.5\Delta \leq x < -0.5\Delta$
100	0	$-0.5\Delta \leq x < 0.5\Delta$
101	Δ	$0.5\Delta \leq x < 1.5\Delta$
110	2Δ	$1.5\Delta \leq x < 2.5\Delta$
111	3Δ	$2.5\Delta \leq x < 3.5\Delta$

Lượng tử đều

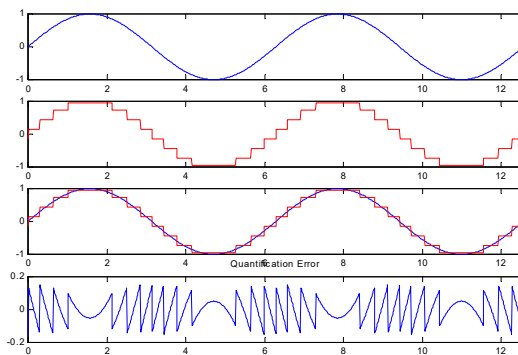


Lượng tử đều

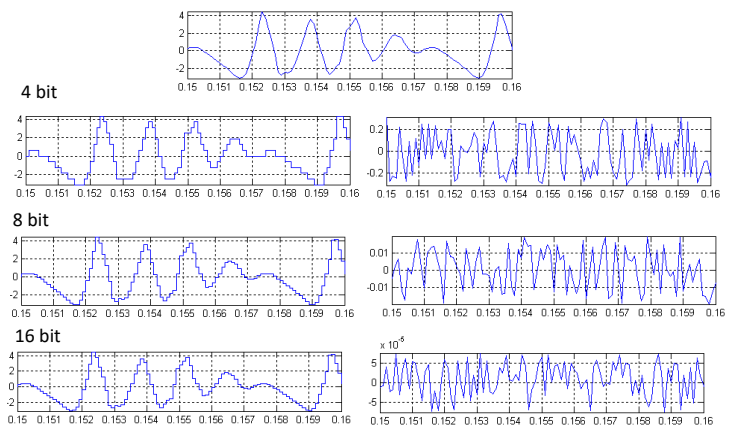
• $L = 16$



Lượng tử đều



Lượng tử đều



Các tính chất lượng tử đều

- Mật độ xác suất sai số lượng tử

$$p_e(\xi) = \sum_{i=-\ell}^{\ell} p_x(i\Delta + \xi), \quad \ell = (L-1)/2$$

phân bố đều giữa $-\Delta/2$ và $+\Delta/2$

$$p_e(\xi) = 1/\Delta, \quad |\xi| \leq \Delta/2 \\ = 0, \quad |\xi| > \Delta/2$$

- Trung bình tạp âm lượng tử = 0
- Phương sai

$$\sigma_e^2 = \int_{-\delta/2}^{\delta/2} \xi^2 / \Delta d\xi = \Delta^2/12$$

Các tính chất lượng tử đều

- Tỷ số tín hiệu trên nhiễu

$$SN = \frac{\sigma_x^2}{\sigma_e^2}$$

$$SN(dB) = 10 \lg \left(\frac{\sigma_x^2}{\sigma_e^2} \right) = 6,02b + 4,77 - 20 \lg \left(\frac{x_{\max}}{\sigma_x} \right)$$

$$\text{Nếu } x_{\max} = 4\sigma_{\max} \rightarrow SN(dB) = 6b - 7,3$$

Với $b \geq 6$, tăng 6 dB mỗi khi tăng 1 bit lượng tử. Để có chất lượng thích hợp cần có $b \geq 11$

- Có thể tính SN như sau:

$$SN = \frac{\frac{1}{N} \sum_{n=0}^{N-1} x^2(n)}{\frac{1}{N} \sum_{n=0}^{N-1} e_q^2(n)} = \frac{\sum_{n=0}^{N-1} x^2(n)}{\sum_{n=0}^{N-1} e_q^2(n)}$$

Tỷ số tín hiệu trên nhiễu

$$\bullet SN = \frac{\text{Năng lượng tín hiệu}}{\text{Năng lượng nhiễu}} = \frac{W_s}{W_n}$$

$$\bullet SN_{dB} = 10 \log_{10} SN$$

Hoặc

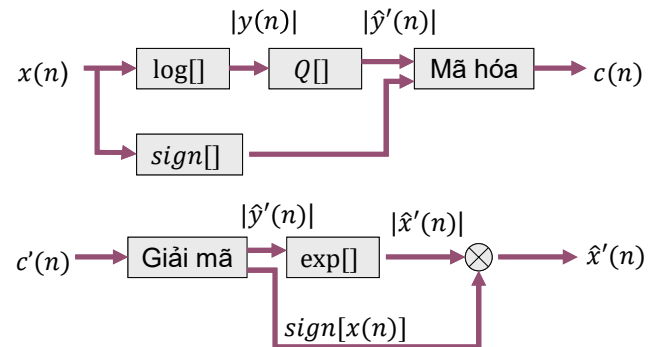
$$SN_{dB} = 20 \log_{10} \frac{\text{Biên độ tín hiệu}}{\text{Biên độ nhiễu}}$$

Tỷ số tín hiệu trên nhiễu

Năng lượng	SN (dB)
Tín hiệu = Nhiễu	0
Tín hiệu = 2 Nhiễu	2
Tín hiệu = 10 Nhiễu	10
Tín hiệu = 100 Nhiễu	20
Tín hiệu = 1000 Nhiễu	30
Tín hiệu = 10^N Nhiễu	N x 10

Lượng tử logarit

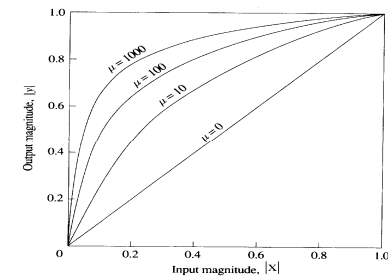
- Sau khi lấy logarit biên độ tín hiệu sẽ mã hoá tuyến tính



Lượng tử logarit

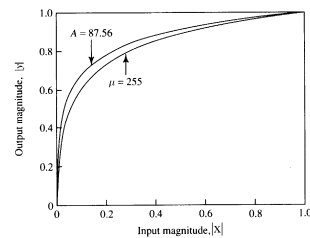
- Hai giải pháp dùng cho điện thoại
 - Luật μ (dùng ở Mỹ)

$$|y| = \frac{\log(1 + \mu|x|)}{\log(1 + \mu)}$$



Lượng tử logarit

- Hai giải pháp dùng cho điện thoại
 - Luật A (dùng ở châu Âu)

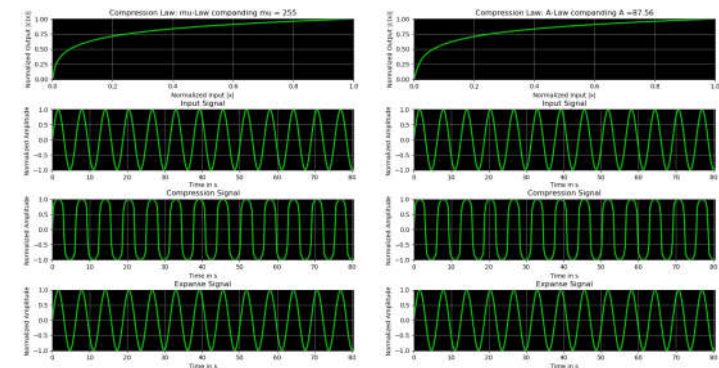


$$|y| = \frac{1 + \log A |x|}{1 + \log A}$$

$$\mu = 255 \sim A = 87,56$$

8 bit logarit ~ 12 bit lượng tử đều

Ví dụ mu-law, a-law



Algorithms

For a given signal x , the output of the μ -law compressor is

$$y = \frac{V \log(1 + \mu|x|/V)}{\log(1 + \mu)} \operatorname{sgn}(x)$$

where V is the maximum value of the signal x , μ is the μ -law parameter of the compander, \log is the natural logarithm, and sgn is the signum function.

The output of the A-law compressor is

$$y = \begin{cases} \frac{A|x|}{1 + \log A} \operatorname{sgn}(x) & \text{for } 0 \leq |x| \leq \frac{V}{A} \\ \frac{V(1 + \log(A|x|/V))}{1 + \log A} \operatorname{sgn}(x) & \text{for } \frac{V}{A} < |x| \leq V \end{cases}$$

where A is the A-law parameter of the compander and the other elements are as in the μ -law case.

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

86

86

Lập trình Python

ONE LOVE. ONE FUTURE.

```

mirror_mod.mirror_object =
operation == "MIRROR_X":
mirror_mod.use_x = True
mirror_mod.use_y = False
mirror_mod.use_z = False
operation == "MIRROR_Y":
mirror_mod.use_x = False
mirror_mod.use_y = True
mirror_mod.use_z = False
operation == "MIRROR_Z":
mirror_mod.use_x = False
mirror_mod.use_y = False
mirror_mod.use_z = True

selection at the end -add
ob.select= 1
for ob.select=1
context.scene.objects.active
("Selected" + str(modifier
mirror_ob.select = 0
bpy.context.selected_obj
data.objects[one.name].sel
print("please select exactly
--- OPERATOR CLASSES ---

(types.Operator):
X mirror to the selected
object.mirror_mirror_x"
error X"

context):
context.active_object is not

```

87

87

Lượng tử thích nghi

- Bước lượng tử tùy thuộc vào biên độ tín hiệu
- Thích nghi trước

$$y(n) = x(n) G(n)$$

$$\hat{y}(n) = Q[y(n)]$$

$$c(n) = \text{Mã hóa}(\hat{y}(n))$$

$$G(n) = \text{Thích nghi độ k.đại}(c(n))$$

$$\hat{y}'(n) = \text{Giải mã}(c(n))$$

$$\hat{x}'(n) = \frac{\hat{y}'(n)}{G'(n)}$$

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

88

88

Lượng tử thích nghi

- Thích nghi sau

$$y(n) = x(n) G(n)$$

$$\hat{y}(n) = Q[y(n)]$$

$$c(n) = \text{Mã hóa}(\hat{y}(n))$$

$$G(n) = \text{Thích nghi độ k.đại}(c(n))$$

$$\hat{y}'(n) = \text{Giải mã}(c(n))$$

$$\hat{x}'(n) = \frac{\hat{y}'(n)}{G'(n)}$$

TRƯỜNG ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

89

89

Một số chuẩn mã hoá âm thanh/tiếng nói

- G.721 : ADPCM, 32 kbps, 4bits, 8kHz
- G.722 : ~ADPCM, 48 đến 64 kbps,
- G.723 : ~ADPCM, 24 kbps, 3 bits, 8kHz
- G.728 : 16 Kbps
- GSM : điện thoại di động, 13 kbps
- Linear Predictive Encoding (Xerox), 5 kbps
- Code Excited Linear Prediction (CELP)
- Digital Video Interactive : ~ADPCM, 4 đến 8 bits
- VoIP: G723.1 (6.4kbits/s), G728, G729 (8kbits/s)