TRƯỜNG ĐẠI HỌC XÂY DỰNG HÀ NỘI
KHOA CÔNG NGHỆ THÔNG TIN

LẬP TRÌNH ÂM THANH

CHƯƠNG 10

Trộn nhiều lớp và âm thanh không gian

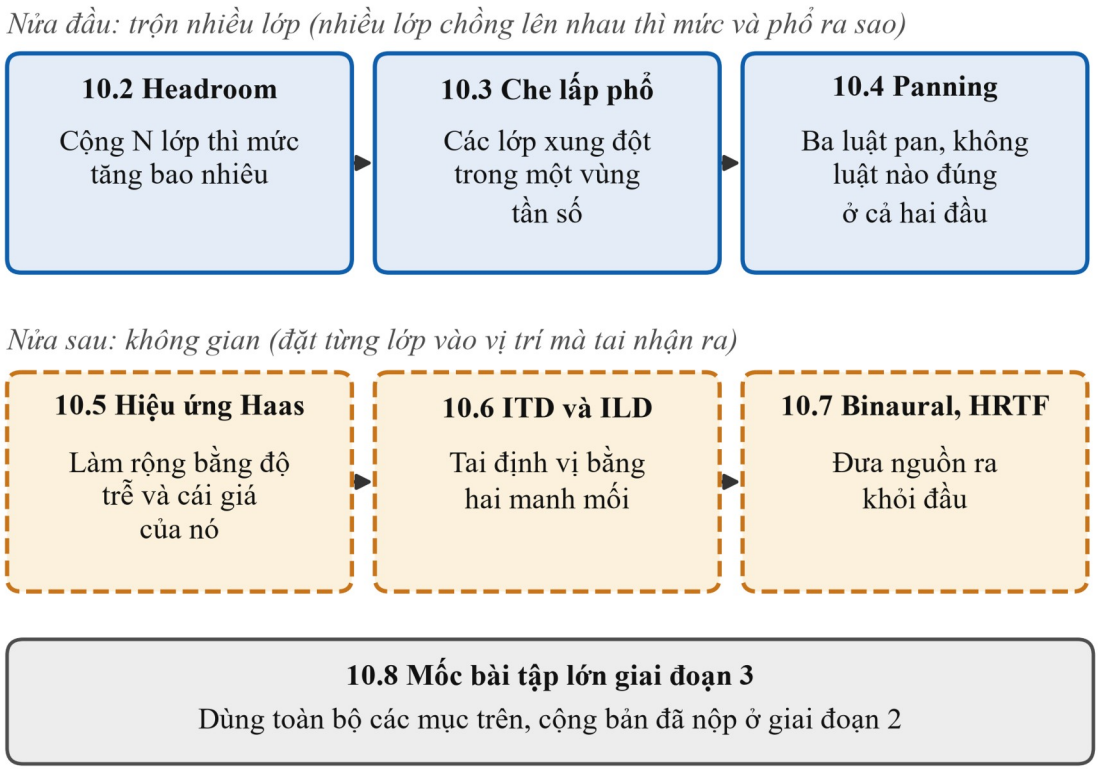
Sau chương này người học phải làm được

- Tính headroom cho N lớp và giải thích vì sao có hai công thức.
- Dùng hệ số tương quan để dự đoán mức tăng và phát hiện lớp bị đảo cực.
- Đo độ che lấp phổ giữa hai lớp, xếp thứ tự việc phải làm theo con số ấy.
- Chọn một trong ba cách tách lớp và nói rõ cái giá phải trả.
- Cài đặt ba luật pan; chứng minh không luật nào đúng ở cả hai đầu.
- Dùng hiệu ứng Haas và đo cái giá của nó khi cộng về mono.
- Tính ITD (Interaural Time Difference – chênh lệch thời gian giữa hai tai), ILD (Interaural Level Difference – chênh lệch mức giữa hai tai) cho một hướng bất kỳ; giải thích thuyết song công bằng đại lượng ka.
- Dựng mô hình binaural từ hai khối của chương 9, đo lại so với công thức.
- Nêu bốn giới hạn của binaural; chọn kích thước khối cho bộ máy thời gian thực.

MỤC 10.1

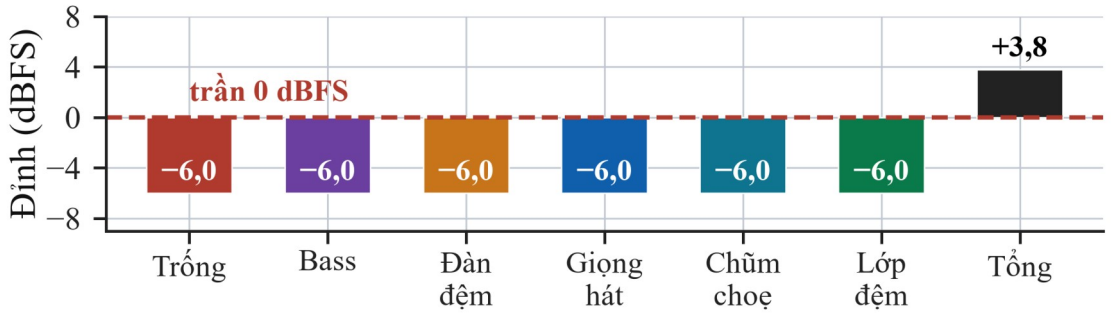
Vị trí chương học trong học phần

- Chín chương trước chỉ nhận vào một tín hiệu và trả ra một tín hiệu.
- Sản phẩm thật có nhiều lớp phát cùng lúc, và người nghe có hai tai.
- Chương này bổ sung hai điều: cộng nhiều lớp cần chừa bao nhiêu, và đặt mỗi lớp vào đâu giữa hai tai.



Từ nhiều lớp rời tới một bản trộn nghe được

- Cho N lớp đã xử lý xong bằng công cụ của các chương trước.
- Chọn gain và vị trí cho từng lớp.
- Bản trộn không cắt đỉnh, các lớp không lấn nhau.
- Bản trộn vẫn nghe được khi cộng về mono.
- Ba yêu cầu mâu thuẫn từng đôi một: cần đo mức mâu thuẫn để chọn có căn cứ.



(a) Sáu lớp, mỗi lớp −6 dBFS: tổng vượt trần

Câu hỏi 1: mức

Tổng vượt trần 3,8 dB. Hạ bao nhiêu, theo công thức nào?

Câu hỏi 2: chỗ ngồi

Các lớp chồng lên nhau thì nghe mờ đi. Làm sao mỗi lớp có chỗ riêng mà không hạ mức?

Chương này trả lời cả hai; câu hỏi 2 có ba cách trả lời

Tần số

Tách bằng EQ (mục 10.3)

Không gian

Tách bằng pan (mục 10.4 – 10.5)

Chiều sâu

Tách bằng vang (chương 9, mục 9.7)

(b) Hai câu hỏi của cả chương

Hình 10.2 — Từ nhiều lớp rời tới một bản trộn nghe được

Bảy mục kỹ thuật và thứ mỗi mục dựng trên

Mục	Nội dung	Dựng trên
10.2	Cộng nhiều lớp và headroom	Chương 5; bộ giới hạn mục 8.4
10.3	Che lấp phổ, ba cách tách lớp	Chương 4; EQ mục 9.4; vang mục 9.7
10.4	Panning và ba luật pan	Hoàn toàn mới
10.5	Hiệu ứng Haas	Trễ mục 6.4; bộ lọc lược mục 8.5
10.6	ITD, ILD và thuyết song công	Trễ phân số mục 9.2; mục 9.3
10.7	Binaural và HRTF	Tích chập theo khối mục 9.8
10.8	Mốc bài tập lớn giai đoạn 3	Bảy mục trên và bản giai đoạn 2

MỤC 10.2

Trộn nhiều lớp và headroom

- Trộn là phép cộng có trọng số: mỗi lớp nhân một gain rồi cộng lại.
- Khó khăn nằm ở việc chọn gain, vì vượt trần là cắt đỉnh chứ không phải to lên.
- Câu hỏi của mục: cộng N lớp thì cần hạ mỗi lớp bao nhiêu.

Công thức trộn

$$y[n] = \sum_{k=1}^{N} g_k x_k[n]$$

(10.1)

- xk là lớp thứ k; gk là gain tính theo hệ số nhân, không theo decibel.
- N là số lớp.
- Cộng N tín hiệu đầy biên độ thì kết quả vượt trần.

Cộng nhiều lớp, mỗi lớp một gain riêng

```
def cong_lop(lop, gain_db=None):
    """Cộng nhiều lớp, mỗi lớp có gain riêng tính bằng decibel."""
    n = max(len(sang_mono(x)) for x in lop)
    stereo = any(np.asarray(x).ndim == 2 for x in lop)
    ra = np.zeros((n, 2)) if stereo else np.zeros(n)
    for i, x in enumerate(lop):
        g = 1.0 if gain_db is None else db_to_lin(gain_db[i])
        v = sang_stereo(x) if stereo else sang_mono(x)
        ra[:len(v)] += g * v          # tự đệm cho bằng lớp dài nhất
    return ra
```

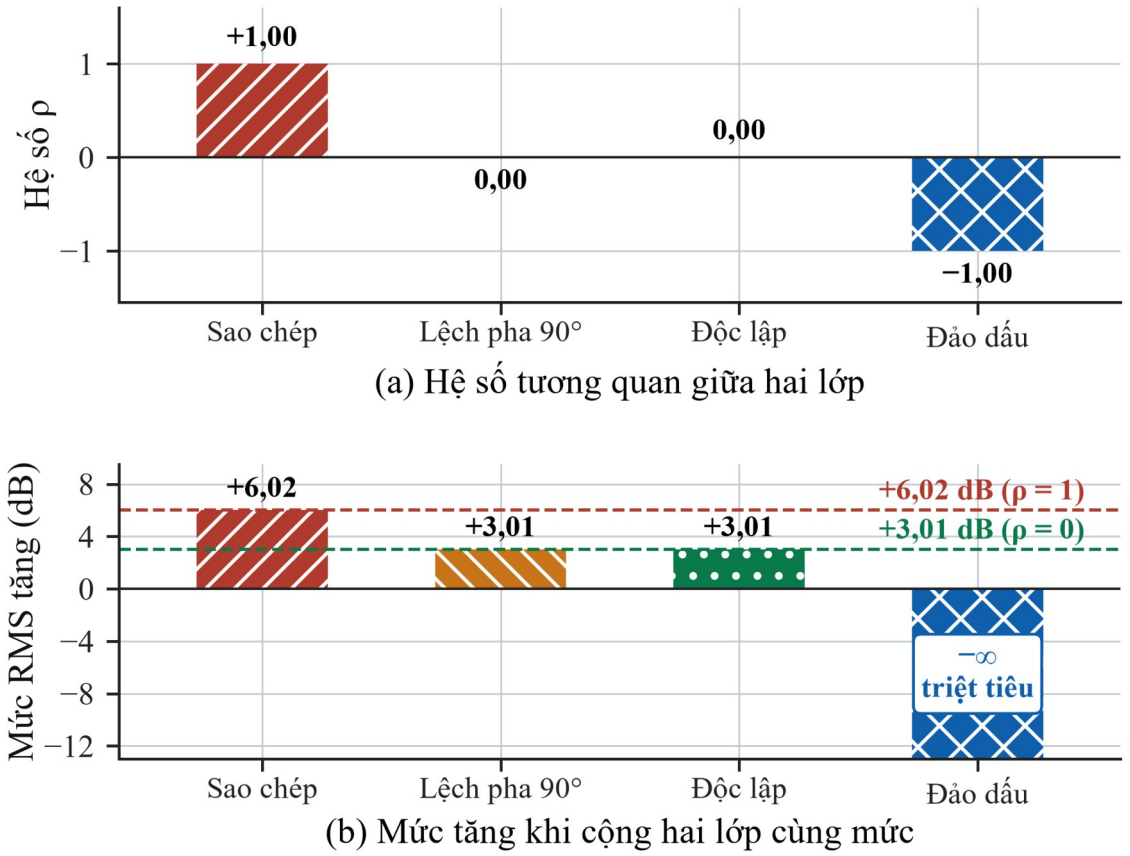
Mức tăng phụ thuộc hệ số tương quan

$$\rho = \frac{\sum_n (a[n] - \bar{a})(b[n] - \bar{b})}{\sqrt{\sum_n (a[n] - \bar{a})^2 \sum_n (b[n] - \bar{b})^2}}$$

(10.2)

- Cộng hai tín hiệu cùng mức: mức tăng là một khoảng chứ không phải một số.
- ρ là hệ số tương quan Pearson, chạy từ −1 tới +1.
- ρ = 0: không liên quan; ρ = +1: giống hệt hình dạng; ρ = −1: bản đảo dấu.

Bốn kiểu quan hệ giữa hai lớp và mức tăng tương ứng



Hình 10.3 — Bốn kiểu quan hệ giữa hai lớp và mức tăng tương ứng

Cộng hai lớp cùng mức: bốn trường hợp

Quan hệ	ρ	Mức RMS tăng	Gặp ở đâu
Cùng tín hiệu, nhân đôi	+1,00	+6,02 dB	Nhân bản track cho dày
Cùng nguồn, lệch pha 90 độ	0,00	+3,01 dB	Hai micro đặt lệch nhau
Hai lớp độc lập	≈ 0,00	+3,01 dB	Hai nhạc cụ khác nhau
Bản gốc và bản đảo dấu	−1,00	Triệt tiêu hoàn toàn	Đấu nhầm dây cân bằng

Khi ρ bằng −1, hai lớp không nhỏ đi mà biến mất

Nghe từng lớp riêng thì bình thường, cộng lại thì mất tiếng. Cách phát hiện rẻ nhất: đo hệ số tương quan giữa từng cặp lớp trước khi trộn; cặp có ρ âm sâu là cặp có vấn đề.

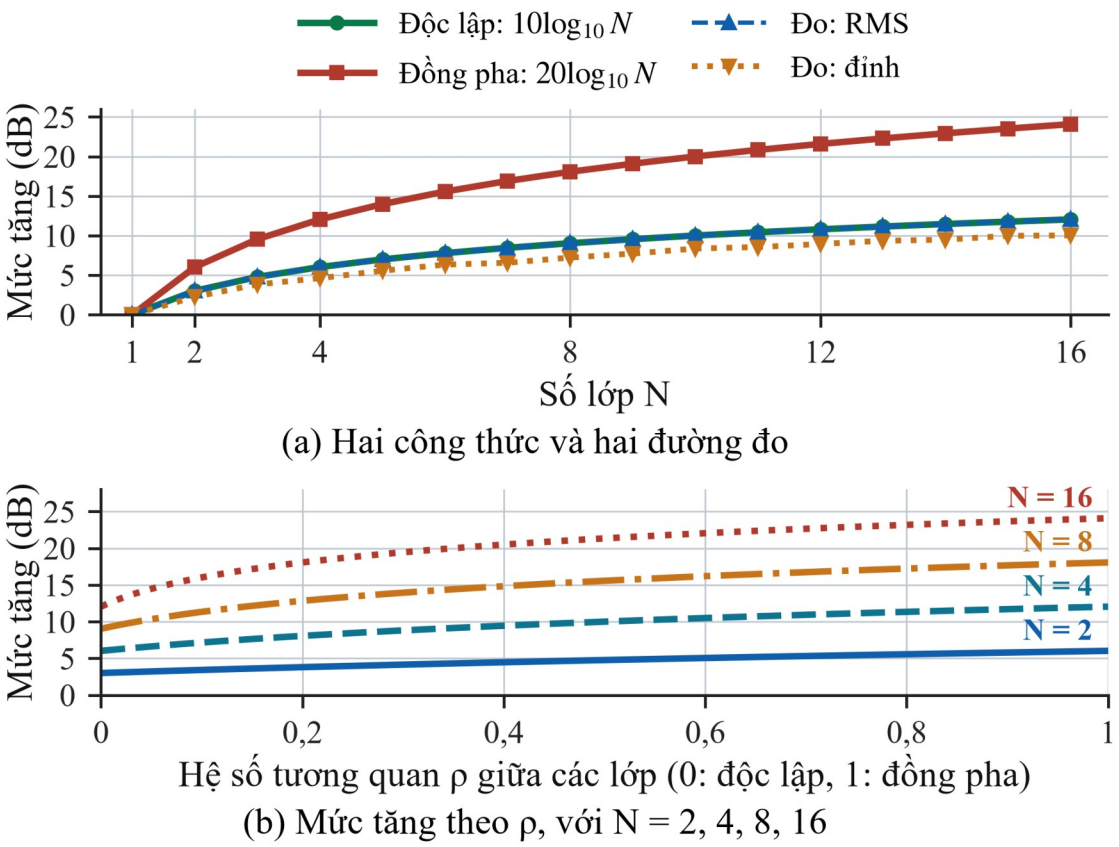
Hai đầu mút của mức tăng theo số lớp

$$\Delta L = 10\log_{10} N \ (\rho = 0)\,; \qquad \Delta L = 20\log_{10} N \ (\rho = 1)$$

(10.3)

- Các lớp độc lập: công suất cộng, mức tăng theo mười lần loga của N.
- Các lớp đồng pha: biên độ cộng, mức tăng gấp đôi con số trên.
- Hạ mỗi lớp đúng lượng ấy trước khi cộng: đó là headroom cần chừa.

Hai công thức là hai đầu mút, chỗ thực tế nằm giữa



Hình 10.4 — Hai công thức là hai đầu mút, và chỗ thực tế nằm giữa

Điều đáng chú ý: đường đỉnh thấp hơn đường RMS

- RMS (Root Mean Square – giá trị hiệu dụng) đo được bám sát đường lý thuyết độc lập: công thức đúng.
- Tám lớp nhiễu độc lập: RMS tăng 8,9 dB, còn đỉnh chỉ tăng 7,2 dB.
- Đỉnh của các lớp hiếm khi rơi đúng vào cùng một mẫu.
- Chừa theo công thức là an toàn nhưng thừa chừng một tới hai decibel.
- Chỗ thừa là dư địa để bộ giới hạn mục 8.4 làm việc.

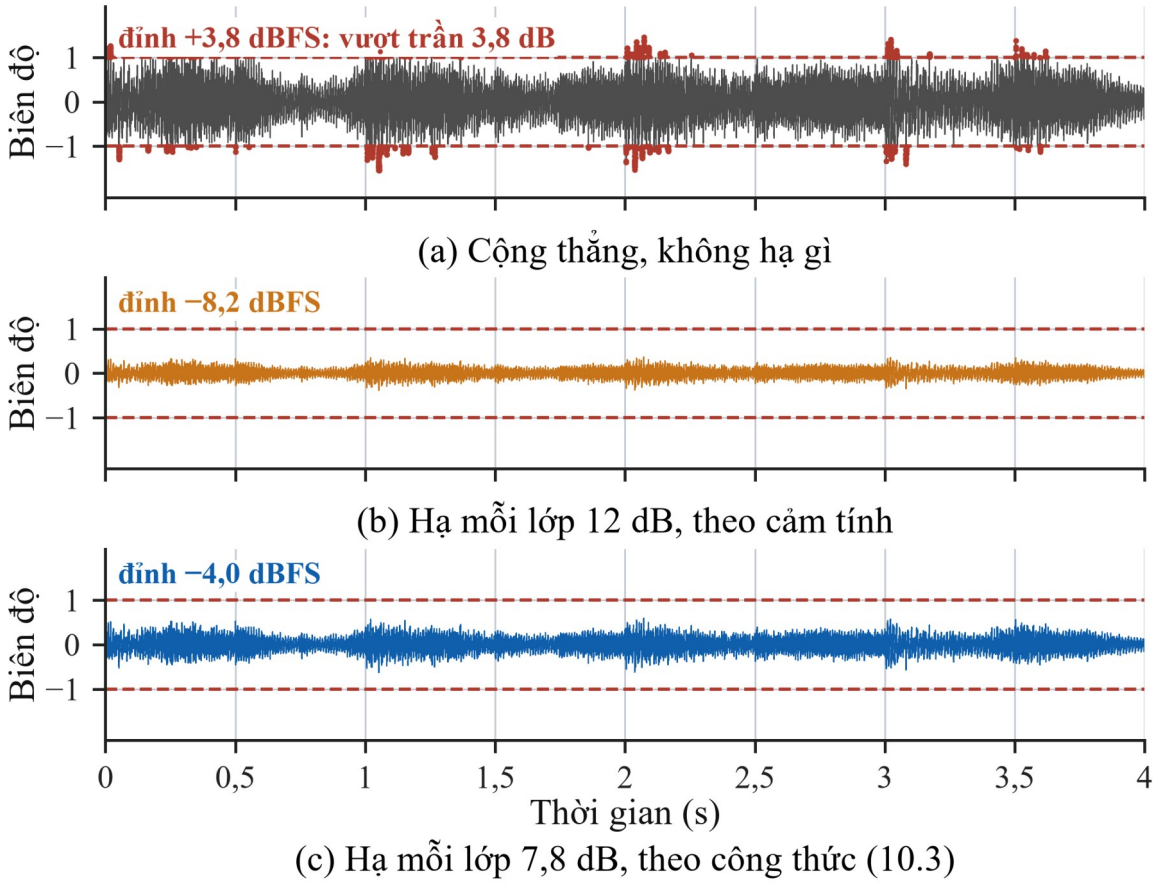
Headroom cần chừa cho từng lớp

Số lớp N	Nếu độc lập	Nếu đồng pha	Ghi chú
1	0,0 dB	0,0 dB	Không phải chừa gì
2	−3,0 dB	−6,0 dB	Chênh lệch còn nhỏ
4	−6,0 dB	−12,0 dB	Mức tối thiểu của bài tập lớn giai đoạn 3
8	−9,0 dB	−18,1 dB	Bản trộn nhỏ điển hình
12	−10,8 dB	−21,6 dB	Gấp đôi số lớp chỉ thêm 3 dB
16	−12,0 dB	−24,1 dB	Nên gom nhóm rồi trộn nhóm

Mười lần loga tăng chậm: thứ đắt là bốn lớp đầu tiên, không phải các lớp thêm sau.

Sáu lớp: cộng thẳng, hạ theo cảm tính, hạ theo công thức

- Sáu lớp: trống, bass, đàn đệm, giọng hát, chũm choẹ, lớp đệm.
- Mỗi lớp đã chuẩn hoá về −6 dBFS (decibel so với mức toàn thang, Full Scale), như người mới trộn hay làm.
- Sáu lớp dùng suốt chương.



Hình 10.6 — Sáu lớp: cộng thẳng, hạ theo cảm tính, và hạ theo công thức

Ba cách xử lý cùng sáu lớp

Cách làm	Kết quả	Nhận xét
Cộng thẳng	Vượt trần, cắt đỉnh nhiều đoạn	Kéo sáu track vào phần mềm rồi phát
Hạ 12 dB mỗi lớp	Không cắt, bản trộn nhỏ hẳn	Phí gần bốn decibel; nâng lại thì nâng cả nhiễu nền
Hạ 7,8 dB theo (10.3)	Không cắt, còn dư địa vừa phải	Mười lần loga của sáu bằng 7,78 dB

Chừa headroom trước, đo khi con số lệch

- **Chừa trước, đừng sửa sau:** hạ từng lớp trước khi cộng để giữ dải động; phần đã cắt đỉnh không lấy lại được.
- **Đo tương quan khi con số không khớp:** đỉnh cao hơn dự đoán nhiều nghĩa là các lớp đang tương quan.
- **Đừng chừa thừa:** mỗi decibel thừa là một decibel phải nâng lại, kéo theo nhiễu nền mục 7.4.

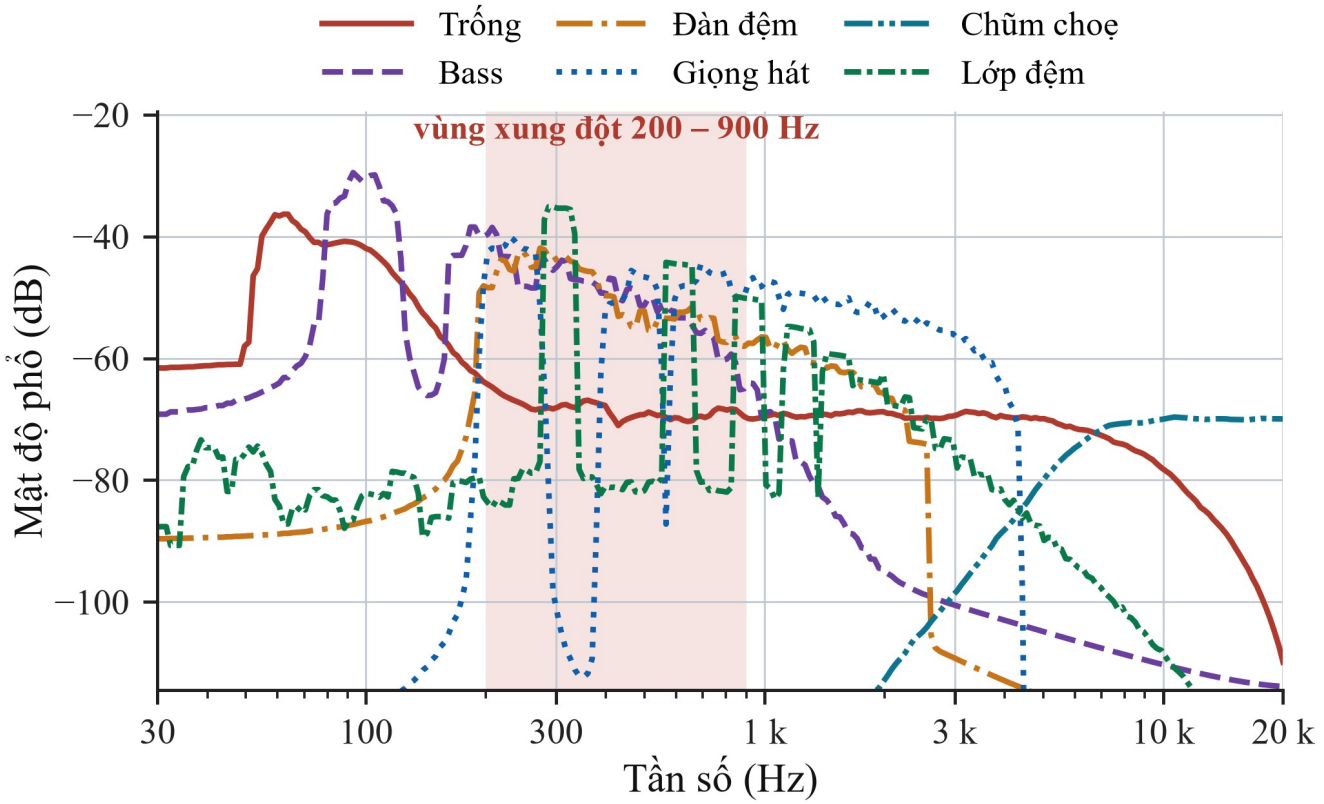
MỤC 10.3

Che lấp phổ và ba cách tách lớp

- Đủ headroom rồi mà bản trộn vẫn nghe đục, khó phân biệt các lớp.
- Nguyên nhân: các lớp chiếm cùng một vùng tần số.
- Câu hỏi của mục: đo mức che lấp bằng số và chọn cách tách có căn cứ.

Phổ của sáu lớp vẽ chồng lên nhau

- Tai phân giải theo dải: lớp mạnh hơn che lớp yếu hơn trong cùng dải.
- Trống và bass nằm dưới 200 Hz; chũm choẹ nằm trên 6 kHz.
- Đàn đệm, giọng hát, lớp đệm chen nhau trong vùng 200 tới 900 Hz.
- Vùng ấy là chỗ bản trộn nghe đục.



Hình 10.7 — Phổ của sáu lớp vẽ chồng lên nhau

Độ che lấp phổ giữa hai lớp

$$M = \sum_{k=1}^{K} \min\left(p_k^A, p_k^B\right), \qquad p_k = \frac{E_k}{\sum_j E_j}$$

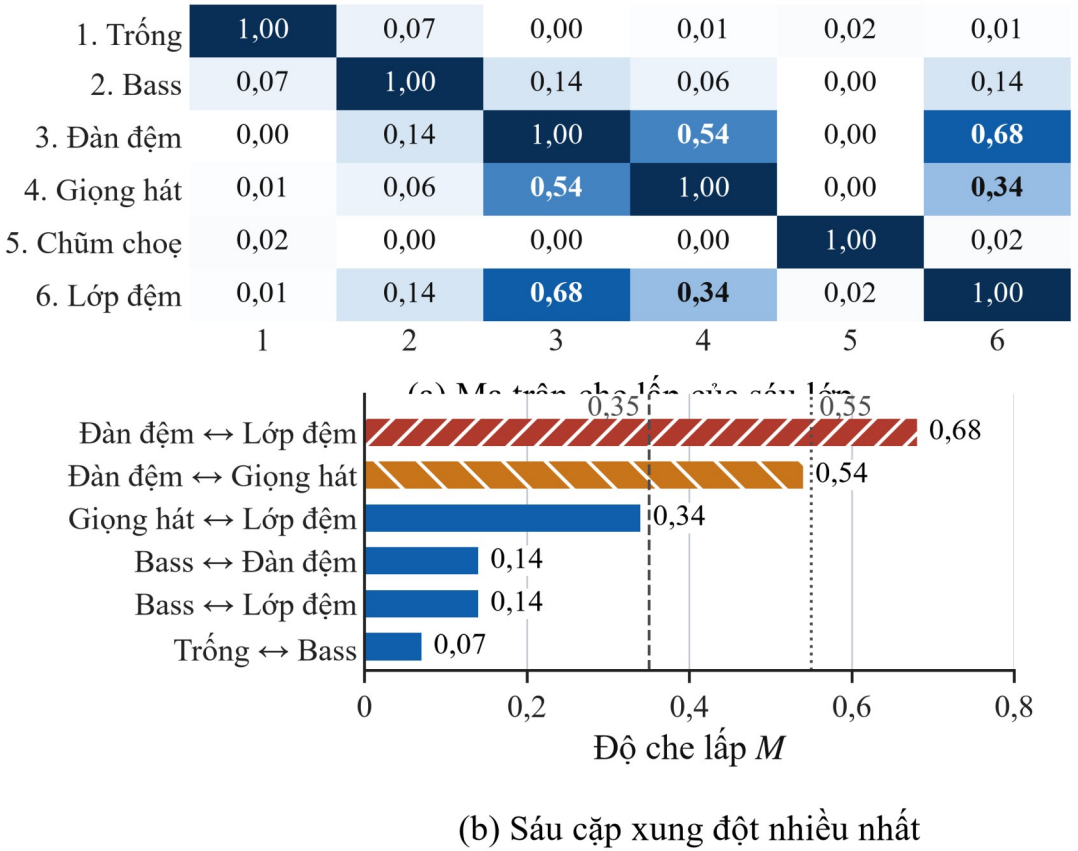
(10.4)

- Chia phổ thành 24 dải theo thang loga, tính tỉ lệ năng lượng mỗi lớp trong từng dải.
- pk là tỉ lệ năng lượng ở dải thứ k; tổng của chúng bằng 1.
- M chạy từ 0 (không đụng nhau) tới 1 (cùng một vùng phổ).

Đo độ che lấp giữa hai lớp

```
def che_lap_pho(a, b, sr=SR, n_dai=24):
    def pho(v):
        f, P = _welch_don_gian(sang_mono(v), sr)
        canh = np.logspace(math.log10(30.0),
                           math.log10(sr / 2 * 0.95), n_dai + 1)
        e = np.array([P[(f >= canh[i]) & (f < canh[i + 1])].sum()
                      for i in range(n_dai)])
        return e / e.sum()        # chuẩn hoá về tỉ lệ
    return float(np.minimum(pho(a), pho(b)).sum())
```

Ma trận che lấp của sáu lớp và sáu cặp xung đột nhiều nhất



(b) Sáu cặp xung đột nhiều nhất

Hình 10.8 — Ma trận che lấp của sáu lớp và sáu cặp xung đột nhiều nhất

Ba cặp xung đột nhiều nhất

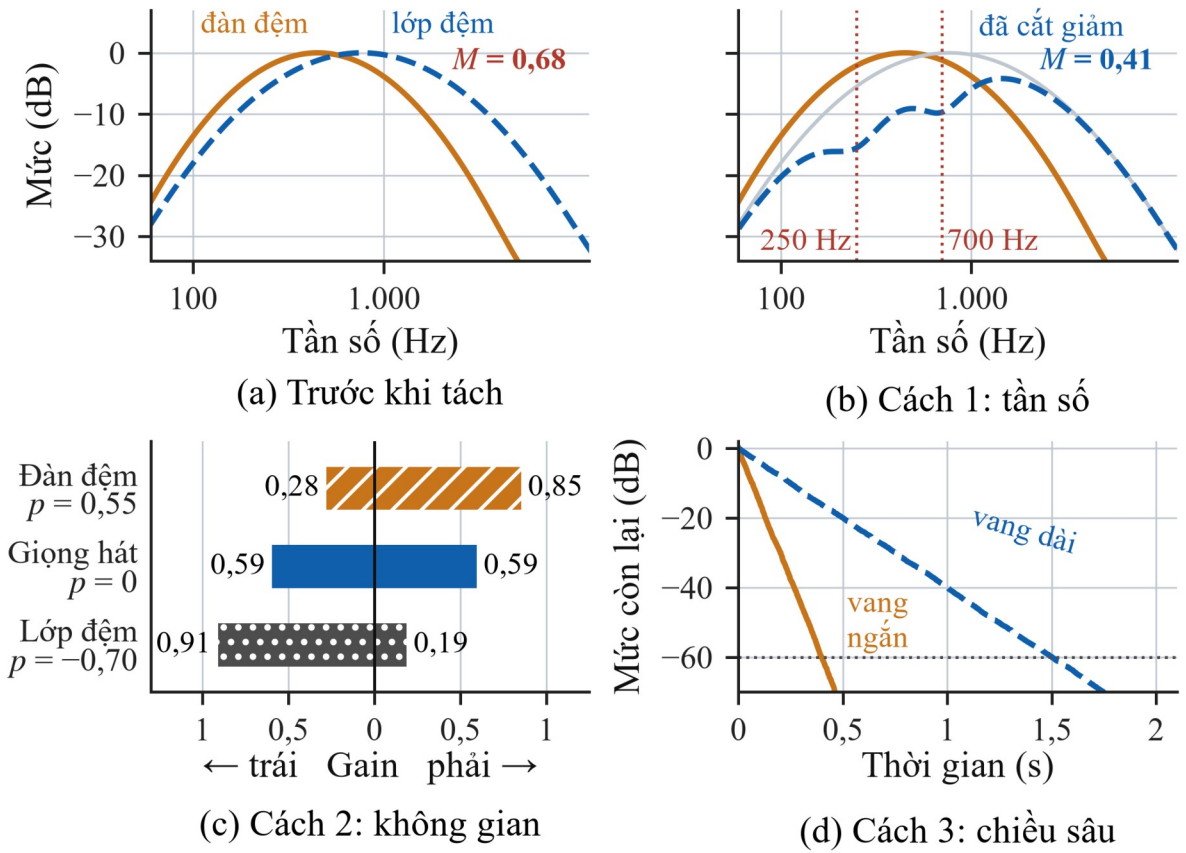
Cặp lớp	Độ che lấp	Đọc ra sao
Đàn đệm ↔ Lớp đệm	0,68	Bất lợi nhất; làm trước tiên
Đàn đệm ↔ Giọng hát	0,54	Đáng lo vì giọng hát không được bị che
Giọng hát ↔ Lớp đệm	0,34	Chấp nhận được; xử lý sau cùng

Ba ngưỡng đọc độ che lấp

- **Dưới 0,35:** hai lớp tự tách, không phải làm gì.
- **Từ 0,35 tới 0,55:** nghe kỹ mới thấy; xử lý nếu một lớp là lớp quan trọng.
- **Trên 0,55:** chắc chắn nghe đục; phải tách.

Ngưỡng rút ra từ cách đo 24 dải; đổi số dải thì con số đổi, không mang sang cách đo khác.

Tách bằng tần số, bằng không gian và bằng chiều sâu



Hình 10.9 — Tách bằng tần số, bằng không gian và bằng chiều sâu

Ba cách tách và cái giá của từng cách

Cách	Công cụ	Được gì	Mất gì
1 · Tần số	EQ mục 9.4	Hiệu quả nhất, đo được ngay	Đổi màu âm lớp bị cắt
2 · Không gian	Pan 10.4, Haas 10.5	Không đụng tới phổ	Lớp không được ở giữa; có thể sai khi về mono
3 · Chiều sâu	Vang mục 9.7	Tách được dù phổ chồng	Thêm đuôi vang, tính lại headroom

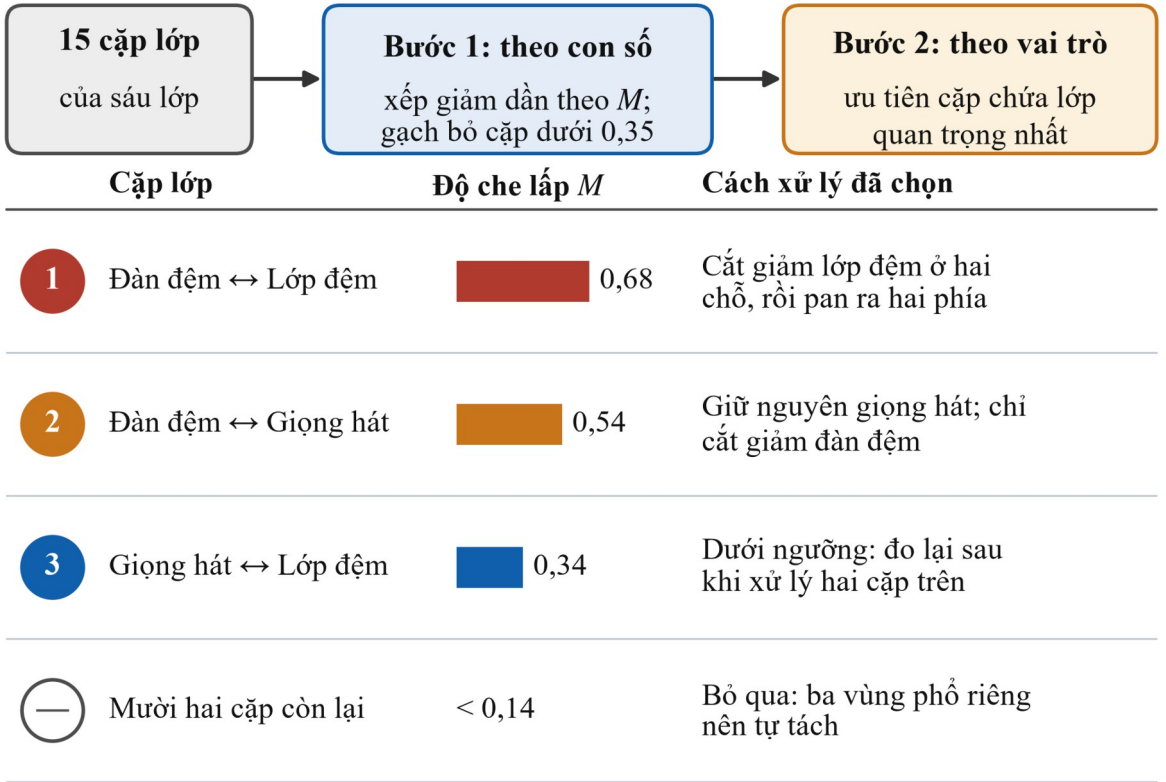
Ba cách cộng dồn được. Ví dụ: cắt lớp đệm 250 Hz hạ 9 dB (Q 0,9) và 700 Hz hạ 8 dB (Q 1,1): 0,68 xuống 0,41.

Cắt giảm lớp này an toàn hơn nâng lớp kia

Hai cách nghe tương đương nhưng khác nhau về headroom: nâng thì đỉnh cao lên và có thể vượt trần, cắt giảm thì đỉnh hạ xuống. Đây là lập luận của mục 9.4.6 áp vào bài toán nhiều lớp.

Xếp thứ tự việc theo độ che lấp và tầm quan trọng

- Sáu lớp có mười lăm cặp; không cần và không thể xử lý hết.
- **Bước một:** xếp theo độ che lấp giảm dần, bỏ mọi cặp dưới 0,35.
- **Bước hai:** ưu tiên cặp chứa lớp quan trọng nhất, thường là giọng hát.
- Con số xếp thứ tự việc; vai trò của lớp quyết định việc nào làm trước.



Hình 10.10 — Xếp thứ tự việc phải làm theo độ che lấp và theo tầm quan trọng

Danh sách việc rút ra từ hai bước

Thứ tự	Cặp lớp	Che lấp	Cách xử lý đã chọn
1	Đàn đệm ↔ Lớp đệm	0,68	Cắt lớp đệm hai chỗ, pan hai lớp ra hai phía
2	Đàn đệm ↔ Giọng hát	0,54	Không đụng giọng hát; chỉ cắt đàn đệm ở vùng giọng trội
3	Giọng hát ↔ Lớp đệm	0,34	Dưới ngưỡng; đo lại sau hai cặp trên
—	Mười hai cặp còn lại	dưới 0,14	Bỏ qua: ba vùng phổ riêng nên tự tách

MỤC 10.4

Panning — đặt lớp vào chỗ giữa hai loa

- Pan chia một lớp mono thành hai kênh với hai gain khác nhau.
- Vị trí p chạy từ −1 (trái hết) tới +1 (phải hết); luật pan sinh cặp (gL, gR) từ p.
- Câu hỏi thật: kéo p qua lại thì nguồn có đổi độ to không. Câu hỏi này không có lời giải đúng.

Luật thứ nhất: hai gain cộng bằng 1

$$g_L = 1 - q, \quad g_R = q; \qquad q = \frac{p+1}{2}, \quad p \in [-1, 1]$$

(10.5)

- Ở chính giữa, mỗi kênh nhận 0,5.
- Ưu điểm: cộng về mono cho đúng tín hiệu gốc ở mọi vị trí pan.
- Nhược điểm: nghe qua hai loa, giữa hụt 3,01 dB về công suất.

Luật thứ hai: tổng bình phương hai gain bằng 1

$$g_L = \cos\frac{\pi q}{2}, \quad g_R = \sin\frac{\pi q}{2}; \qquad g_L^2 + g_R^2 = 1$$

(10.6)

- Nghe qua hai loa: độ to phẳng ở mọi vị trí pan.
- Cộng về mono: ở giữa mỗi kênh 0,707, tổng 1,414, vống 3,01 dB.
- Phát trên loa điện thoại, mọi thứ ở giữa to hơn ý định 3 dB.

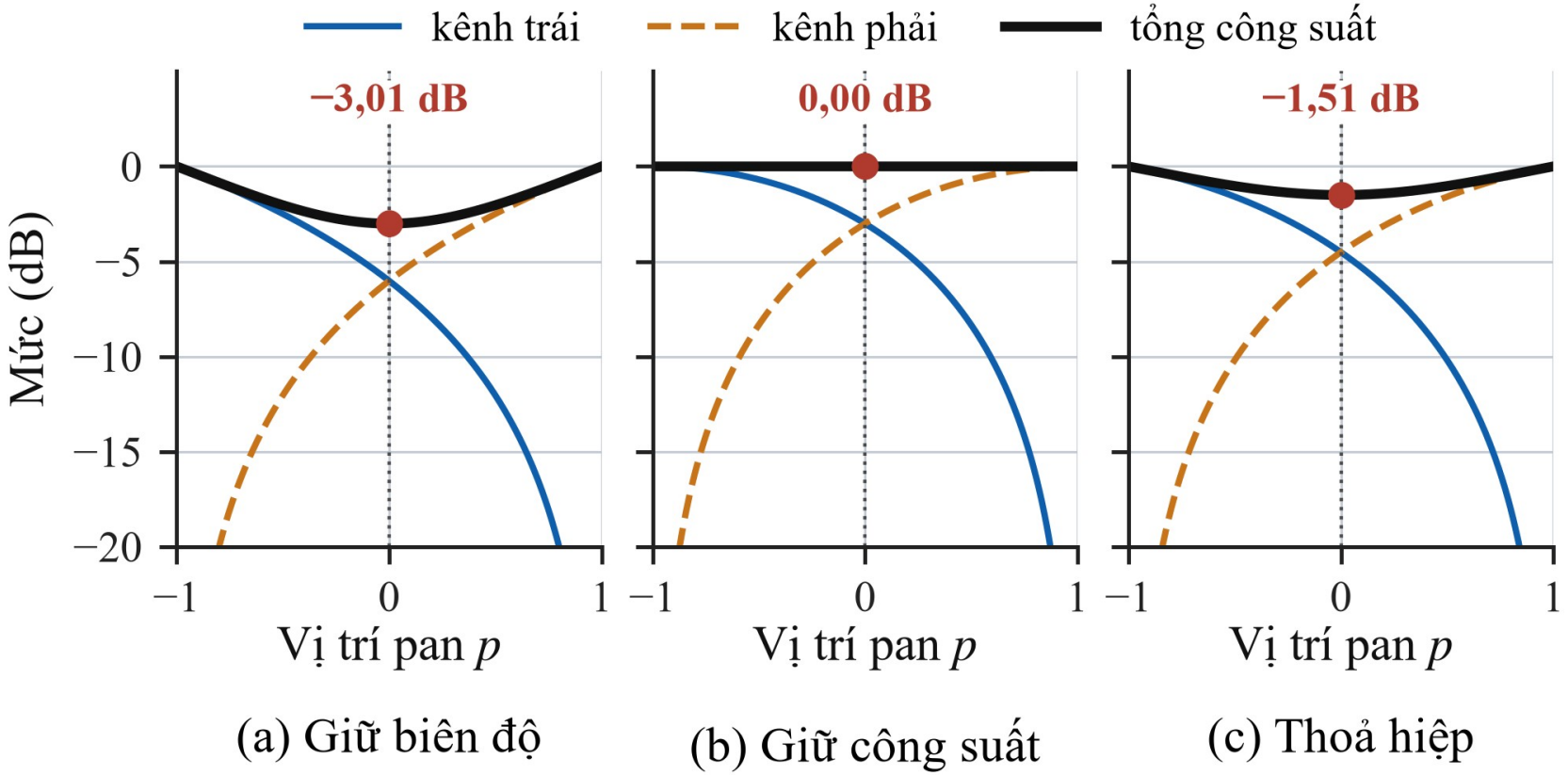
Hai phép đo trên cùng một cặp gain

$$L_{\text{cs}} = 10\log_{10}\left(g_L^2 + g_R^2\right); \qquad L_{\text{mono}} = 20\log_{10}\left(g_L + g_R\right)$$

(10.7)

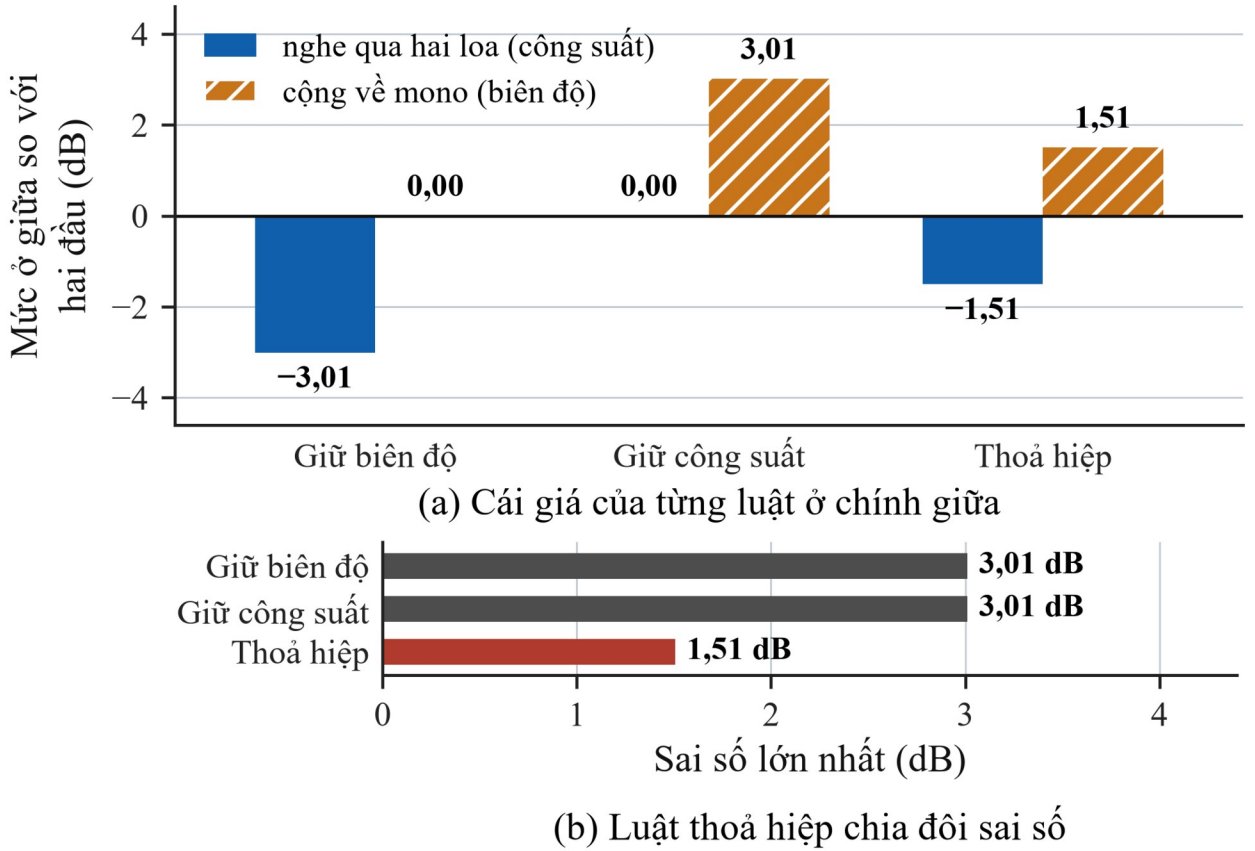
- Dòng trên: mức nghe qua hai loa, công suất cộng lại.
- Dòng dưới: mức sau khi cộng về mono, biên độ cộng lại.
- Không cặp gain nào giữ phẳng cả hai dòng: đó là sự thật toán học, không phải hạn chế thiết bị.

Ba luật pan nhìn theo đường gain



Hình 10.11 — Ba luật pan nhìn theo đường gain

Nghe loa và cộng mono: cái giá ở hai đầu



Hình 10.12 — Cái giá của từng luật ở hai đầu: nghe loa và cộng mono

Ba luật pan và cái giá của mỗi luật

Luật	Hụt ở giữa (loa)	Vống khi cộng mono	Dùng khi
Giữ biên độ	−3,01 dB	0,00 dB	Chắc chắn nghe mono: truyền thanh
Giữ công suất	0,00 dB	+3,01 dB	Chỉ nghe hai loa hoặc tai nghe: trò chơi, phim
Thoả hiệp	−1,51 dB	+1,51 dB	Mặc định của bàn trộn: chia đôi sai số

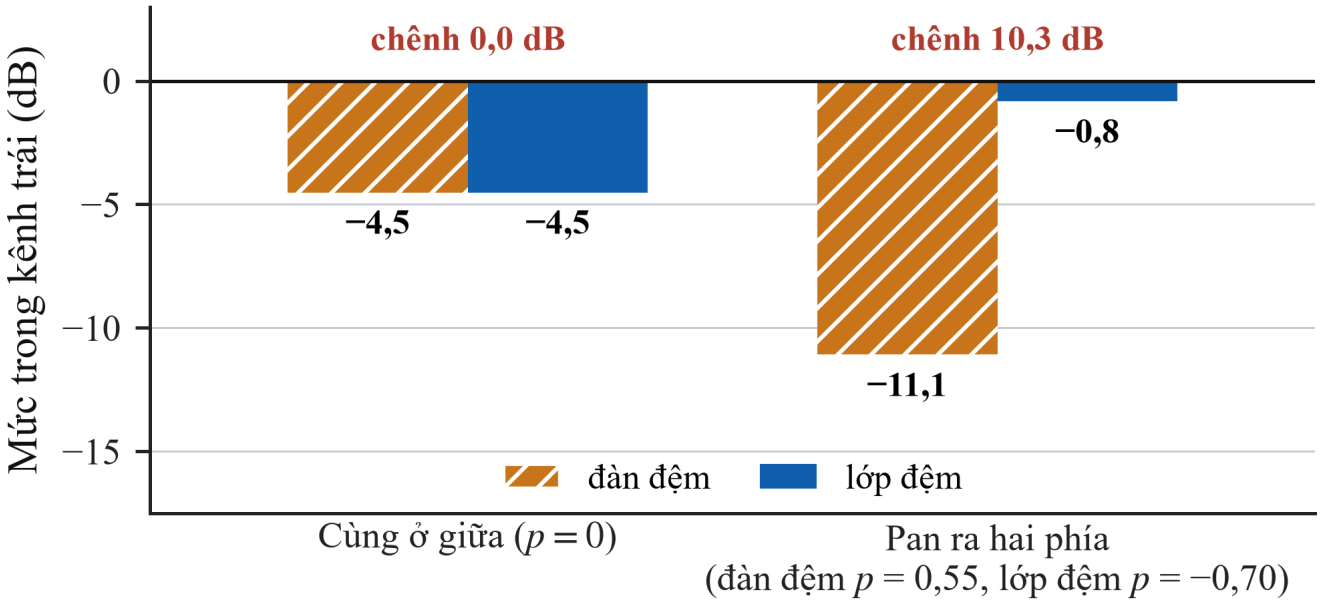
Ba luật pan trong một hàm

```
LUAT_PAN = ('tuyen_tinh', 'cong_suat', 'thoa_hiep')

def luat_pan(p, kieu='thoa_hiep'):
    q = (float(p) + 1.0) / 2.0           # đưa p về [0, 1]
    if kieu == 'tuyen_tinh':
        return 1.0 - q, q
    if kieu == 'cong_suat':
        return math.cos(math.pi * q / 2), math.sin(math.pi * q / 2)
    gl = math.sqrt((1.0 - q) * math.cos(math.pi * q / 2))
    gr = math.sqrt(q * math.sin(math.pi * q / 2))
    return gl, gr                        # trung bình nhân hai luật trên
```

Pan cặp xung đột ra hai phía và đo lại độ tách

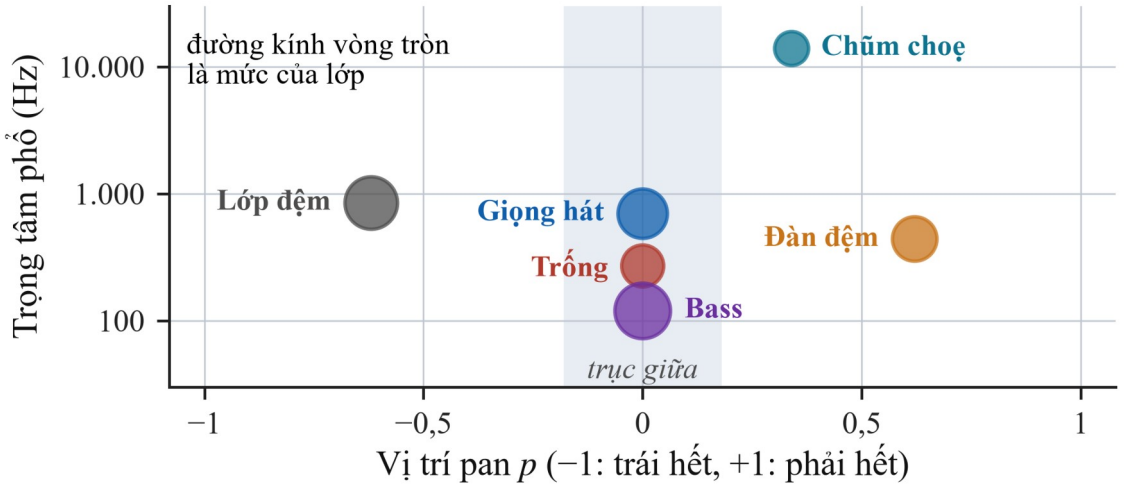
- Cả hai lớp ở giữa: trong mỗi kênh chúng ngang mức, tai không có gì để bám.
- Pan ra hai phía: trong kênh trái, một lớp to hơn lớp kia hơn 10 dB.
- Độ che lấp đo trên tín hiệu mono vẫn nguyên 0,68.
- Phép này tách theo một chiều khác hẳn.



(a) Chênh mức trong một kênh, tai dùng để tách hai lớp

Hình 10.13 — Pan cặp xung đột ra hai phía và đo lại độ tách

Sáu lớp trên bản đồ pan và bốn con số đo lại



(a) Bản đồ đặt chỗ cho sáu lớp

Độ rộng stereo	0,000 → 0,124
Tách cặp bất lợi nhất trong một kênh (dB)	0,0 → 11,4
Gợn phổ khi cộng về mono (dB)	0,00 → 0,32
Đỉnh của bản trộn (dBFS)	−8,47 → −8,03

(b) Bốn con số đo trước và sau khi đặt chỗ

Hình 10.14 — Sáu lớp trên bản đồ pan và bốn con số đo lại

Bốn con số đo trước và sau khi đặt chỗ

Con số	Trước	Sau	Đọc ra sao
Độ rộng stereo	0,000	0,124	Vùng lành mạnh khoảng 0,10 tới 0,60
Tách cặp bất lợi nhất	0,0 dB	11,4 dB	Thứ tai dùng để phân biệt hai lớp
Gợn phổ khi về mono	0,00 dB	0,32 dB	Dưới nửa decibel: an toàn
Đỉnh bản trộn	−8,47 dBFS	−8,03 dBFS	Gần như không đổi

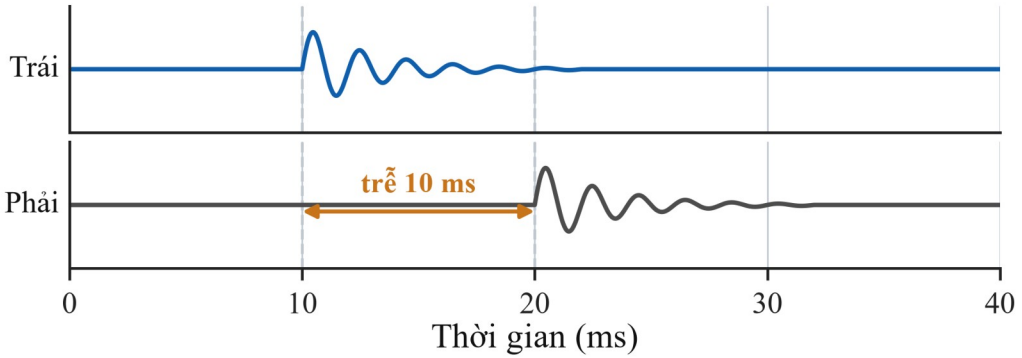
Đỉnh nhích lên 0,44 dB: luật thoả hiệp cho gain gần 0,94 ở một kênh khi pan hết cỡ. Cần đo lại đỉnh sau khi đặt chỗ.

MỤC 10.5

Hiệu ứng Haas — làm rộng bằng thời gian

- Cho một lớp mono vào hai kênh cùng mức, trễ một kênh vài chục mili giây.
- Tai nghe một tiếng, đến từ phía kênh sớm, và rộng hơn hẳn.
- Câu hỏi của mục: trễ bao nhiêu thì dùng được, và cái giá khi cộng mono là gì.

Hai kênh cùng mức, chỉ khác thời điểm, và bảng năm mốc trễ



(a) Hai kênh cùng mức, chỉ khác thời điểm

Khoảng trễ	Tai nghe ra gì	Dùng được không
Dưới 1 ms	Lọc lược, đổi màu âm	Không: đây là lỗi
1 – 5 ms	Nguồn dịch hẳn về kênh sớm	Rộng mạnh, mono kém nhất
5 – 35 ms	Một tiếng, từ phía kênh sớm	Vùng Haas đúng nghĩa
35 – 50 ms	Bắt đầu nghe thành hai sự kiện	Vùng chuyển tiếp
Trên 50 ms	Hai tiếng tách hẳn	Không còn là Haas (slapback)

(b) Năm mốc trễ và kết quả nghe được

Hình 10.15 — Hai kênh cùng mức, chỉ khác thời điểm — và bảng năm mốc trễ

Trễ bao nhiêu thì nghe ra gì

Khoảng trễ	Tai nghe ra gì	Dùng được không
Dưới 1 ms	Lọc lược, đổi màu âm	Không: đây là lỗi
1 – 5 ms	Nguồn dịch hẳn về kênh sớm	Làm rộng mạnh; tương thích mono kém nhất
5 – 35 ms	Một tiếng, từ phía kênh sớm	Vùng Haas đúng nghĩa
35 – 50 ms	Bắt đầu nghe thành hai sự kiện	Vùng chuyển tiếp, tuỳ tư liệu
Trên 50 ms	Hai tiếng tách hẳn	Không còn là Haas: slapback mục 8.7

Cùng một dòng mã, chỉ đổi con số mili giây là ra năm kết quả khác nhau.

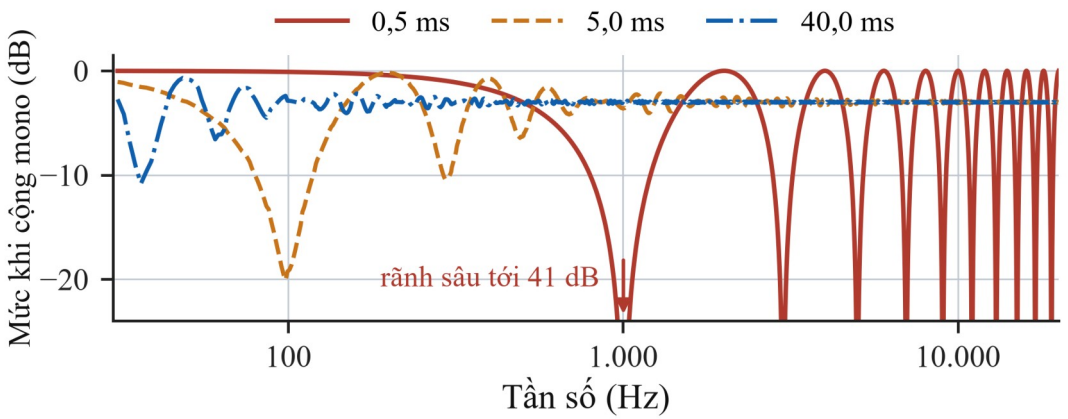
Cộng mono chính là bộ lọc lược

$$m[n] = \frac{x[n] + x[n - D]}{2}; \qquad f_{\text{rãnh}} = \frac{(2k+1)f_s}{2D}$$

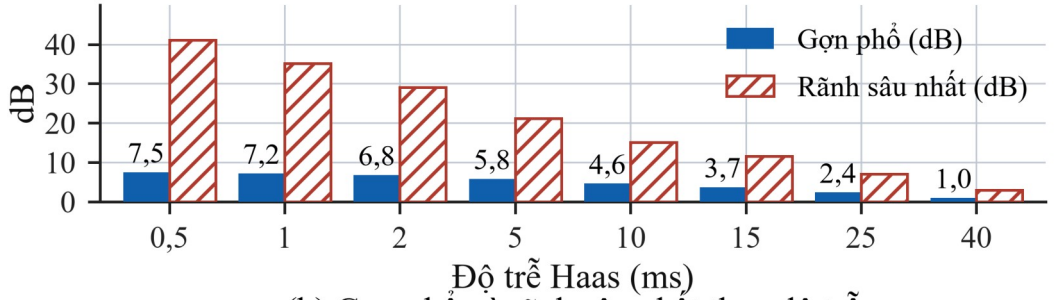
(10.8)

- Dòng thứ nhất: phép cộng về mono.
- Dòng thứ hai: vị trí các rãnh; D là số mẫu trễ, k chạy từ 0.
- Rãnh đầu tiên nằm ở nửa nghịch đảo của thời gian trễ.

Gợn phổ khi cộng về mono, theo lượng trễ



(a) Đáp ứng khi cộng mono (5 và 40 ms làm trơn 1/3 octave)



(b) Gợn phổ và rãnh sâu nhất theo độ trễ

Hình 10.16 — Gợn phổ khi cộng về mono, theo lượng trễ

Trễ và cái giá tương ứng khi cộng về mono

Trễ	Rãnh đầu tiên	Gợn phổ	Nhận xét
0,5 ms	1000 Hz	7,47 dB	Rãnh giữa vùng nghe, sâu tới 40 dB: nguy nhất
5 ms	100 Hz	vài decibel	Rãnh dày, tai coi như đổi màu nhẹ
15 ms	33,3 Hz	nhỏ	Rãnh nằm gọn trong một dải của tai
40 ms	12,5 Hz	1,00 dB	Gần như không đo được

Trễ ngắn mới nguy, trễ dài lại an toàn hơn

Rãnh cách nhau đúng nghịch đảo của thời gian trễ. Trễ ngắn cho rãnh thưa, rơi giữa vùng nghe.
Trễ dài cho rãnh dày, tai gộp thành một mức trung bình phẳng.

Bốn cách làm rộng một lớp, được gì và mất gì

- **Pan:** đổi mức; mono không sao nhưng mất chỗ ở giữa.
- **Haas:** đổi thời điểm; rộng mà gần giữa nhưng về mono thành lọc lược.
- **Hai bản thu:** rộng thật, mono tốt nhất; nhưng phải thu lại.
- **Vang ngắn khác nhau hai bên:** rộng, tự nhiên; nhưng đắt và nghe xa.

Cách	Làm gì	Được	Mất
Pan	Đổi mức hai kênh	Cộng mono không sao; vị trí rõ ràng	Nguồn dịch hẳn sang một bên, mất chỗ ở giữa
Haas	Đổi thời điểm hai kênh	Rất rộng mà vẫn nghe gần giữa	Cộng mono thành lọc lược
Hai bản thu	Thu hai lần rồi pan hai bên	Rộng thật, tương thích mono tốt nhất	Phải thu lại; không dùng cho tư liệu có sẵn
Vang ngắn	Đặt hai lượng vang khác nhau hai bên	Rộng và tự nhiên; không đụng vị trí	Đắt; dễ làm bản trộn nghe xa xăm

Quy tắc chung: kiểm tra mono trước khi giao bản trộn.
Một nút bấm bắt được cả bốn lỗi trên.

Hình 10.17 — Bốn cách làm rộng một lớp, được gì và mất gì

Bấm nút mono và nghe lại

Lớp mỏng đi hoặc biến mất: có Haas trễ ngắn hoặc lớp bị đảo cực. Mọi thứ ở giữa vống lên: đang dùng luật giữ công suất. Đo bằng hàm gon_pho_khi_ve_mono_db: dưới 3 dB là đạt.

MỤC 10.6

ITD và ILD — tai người định vị bằng gì

- Pan và Haas hiệu quả vì mô phỏng hai manh mối thật của tai.
- Chênh lệch thời gian giữa hai tai (ITD) và chênh lệch mức giữa hai tai (ILD).
- Mục này đo cả hai bằng công thức.

Công thức Woodworth

$$\mathrm{ITD}(\theta) = \frac{a}{c}(\theta + \sin\theta)$$

(10.9)

- Sóng từ bên phải chạm tai phải trước, rồi vòng quanh đầu tới tai trái.
- a là bán kính đầu, c là tốc độ âm.
- Dùng a = 8,75 cm và c = 343 m/s.

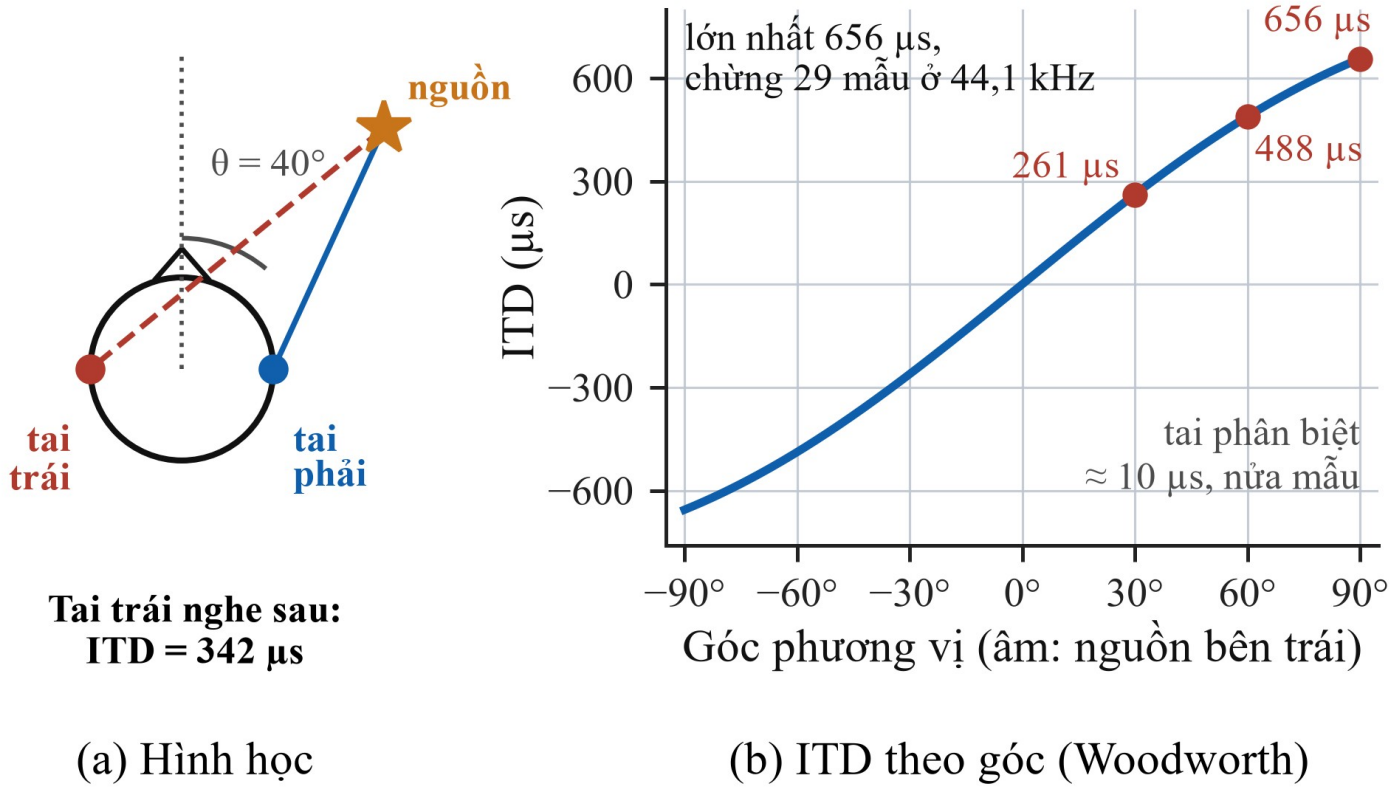
Giá trị lớn nhất khi nguồn ở đúng bên hông

$$\mathrm{ITD}_{\max} = \frac{a}{c}\left(\frac{\pi}{2} + 1\right) \approx 656\,\mu\mathrm{s}$$

(10.10)

- Toàn bộ khả năng định vị trái phải nằm trong khoảng 656 µs.
- Chưa tới ba mươi mẫu ở 44,1 kHz.
- Tai phân biệt được chừng 10 µs, tức nửa mẫu.

ITD theo góc, và độ phân giải của tai



(a) Hình học

(b) ITD theo góc (Woodworth)

Hình 10.18 — ITD theo góc, và độ phân giải của tai

ITD theo Woodworth và phép đi ngược lại

```
BAN_KINH_DAU = 0.0875          # mét — bán kính đầu người lớn
TOC_DO_AM = 343.0              # mét trên giây ở 20 độ C

def itd_giay(goc_do, ban_kinh=BAN_KINH_DAU, c=TOC_DO_AM):
    """0 độ là chính diện, +90 là bên phải. Dương: tai phải trước."""
    th = math.radians(float(goc_do))
    return float(ban_kinh / c * (th + math.sin(th)))

def goc_tu_itd(itd_s, ban_kinh=BAN_KINH_DAU, c=TOC_DO_AM):
    """Woodworth không có nghiệm kín — dò nhị phân."""
    dau = 1.0 if itd_s >= 0 else -1.0
    lo, hi = 0.0, 90.0
    for _ in range(60):
        gi = (lo + hi) / 2
        if itd_giay(gi, ban_kinh, c) < abs(itd_s):
            lo = gi
        else:
            hi = gi
    return float(dau * (lo + hi) / 2)
```

ILD và đại lượng ka

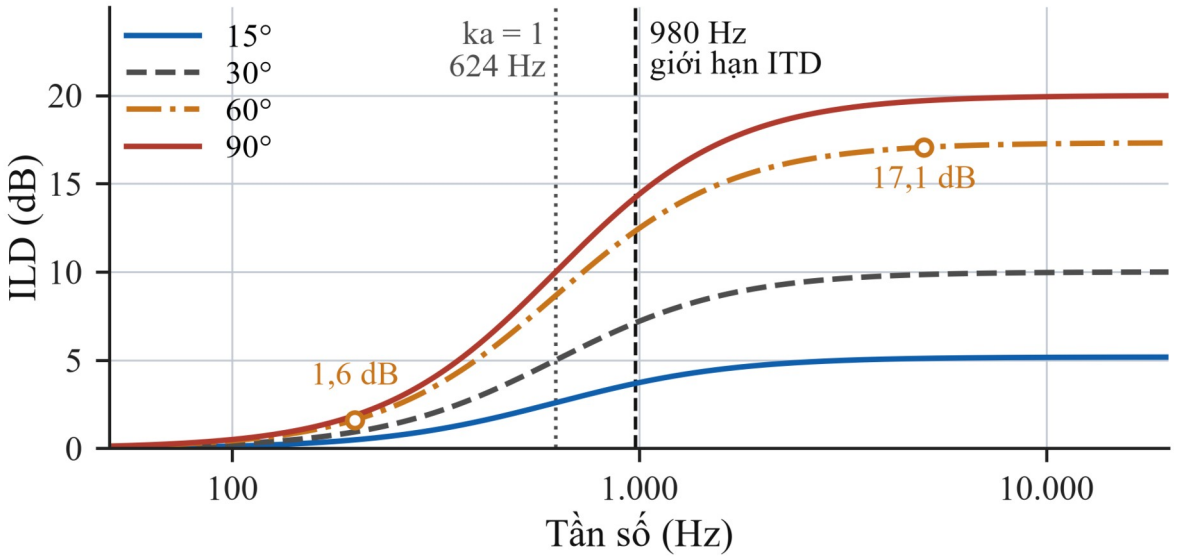
$$ka = \frac{2\pi f a}{c}\ ; \qquad \mathrm{ILD} \approx 20 \frac{(ka)^2}{1+(ka)^2} \sin\theta \ \ [\mathrm{dB}]$$

(10.11)

- ka là tích của số sóng k với bán kính đầu a.
- ka nhỏ hơn 1: sóng vòng qua đầu, chênh mức gần bằng không.
- ka lớn hơn 1: đầu chắn hẳn, số hạng bóng đầu bão hoà về 1.
- Chuyển tiếp ở ka = 1, tức 624 Hz.

ILD theo tần số và theo góc

- Ở góc 60 độ, chênh mức chỉ 1,6 dB tại 200 Hz.
- Cũng góc ấy, chênh mức lên tới 17,1 dB tại 5 kHz.
- Cùng nguồn, cùng hướng, hai dải cho hai con số cách nhau hơn mười lần.



(a) ILD theo tần số, với bốn góc nguồn



(b) Vùng làm việc của từng manh mối

Hình 10.19 — ILD theo tần số và theo góc — thuyết song công

Ranh giới giữa ITD và ILD

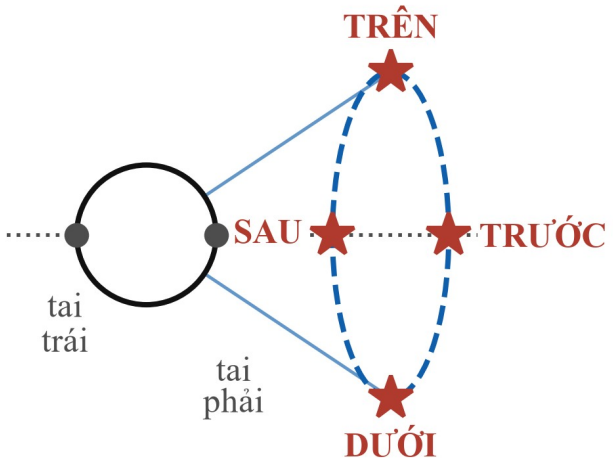
$$f_{\text{gh}} = \frac{c}{4a} \approx 980\ \text{Hz}$$

(10.12)

- Dải trầm định vị bằng ITD, dải cao định vị bằng ILD.
- Trên tần số này nửa bước sóng ngắn hơn khoảng cách hai tai.
- Chênh pha vượt 180 độ nên pha bị lẫn vòng, ITD hết dùng được.
- Đúng lúc ấy ILD đủ lớn: hai manh mối bù cho nhau.

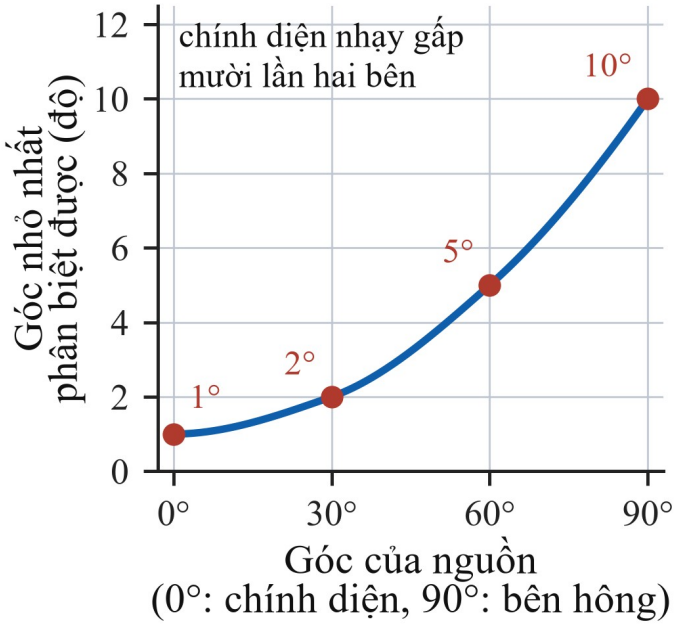
Nón lẫn lộn, và độ nhạy hướng theo góc

- ITD và ILD cho biết trái hay phải, không cho biết trên hay dưới, trước hay sau.
- Mọi hướng trên một mặt nón quanh trục hai tai cho cùng ITD và ILD.
- Góc nhỏ nhất phân biệt được: 1 độ ở chính diện, 10 độ ở bên hông.
- Vì vậy lớp quan trọng nhất được để ở giữa.



Mọi hướng trên mặt nón quanh trục hai tai cho cùng ITD và ILD

(a) Nón lẫn lộn quanh trục hai tai



(b) Độ nhạy theo góc

Hình 10.20 — Nón lẫn lộn, và độ nhạy hướng theo góc

Bốn cách giải quyết nón lẫn lộn

Cách	Cơ chế	Lập trình được không
Dấu vết phổ vành tai	Rãnh phổ đổi theo độ cao	Được, cần HRTF đo thật; mỗi người một khác
Chuyển động đầu	ITD và ILD đổi theo chiều khác nhau trước/sau	Được nếu có cảm biến hướng; hiệu quả nhất
Thông tin thị giác	Mắt thấy nguồn ở đâu, tai tin theo	Được, gần như miễn phí trong trò chơi
Hiểu biết bối cảnh	Bước chân ở dưới, tiếng chim ở trên	Không, nhưng nên biết là có

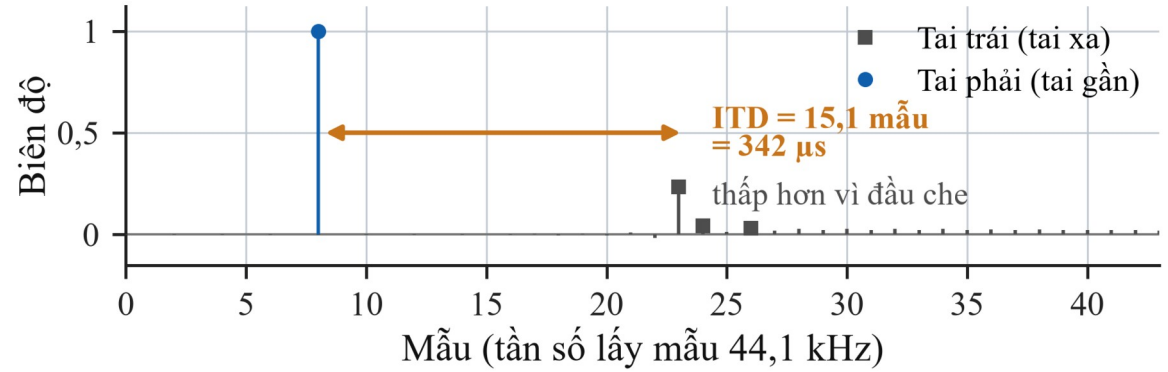
MỤC 10.7

Binaural và HRTF

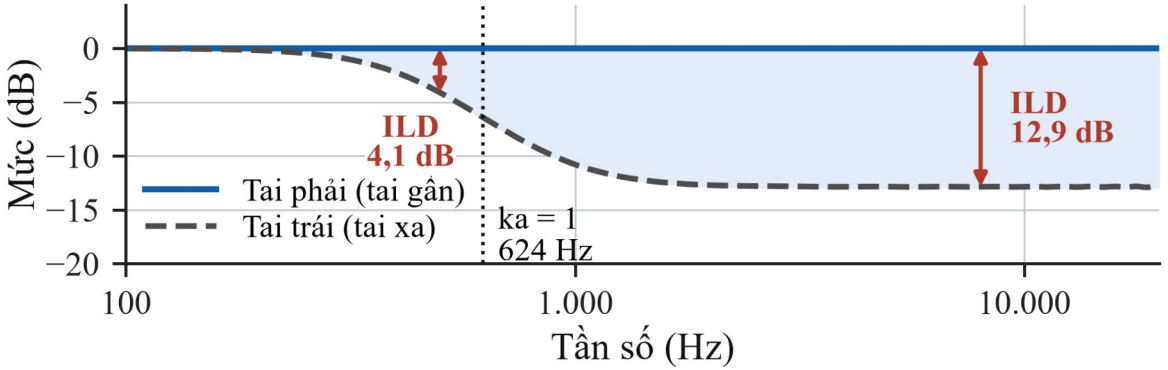
- Đặt nguồn vào một hướng: trễ tai xa đúng ITD, hạ mức tai xa theo ILD.
- Tai thật nhận cả một đáp ứng xung cho mỗi tai, đã bị đầu, vai, vành tai làm biến dạng.
- Cặp đáp ứng xung ấy là HRIR (Head-Related Impulse Response – đáp ứng xung liên quan đến đầu); nhìn theo tần số là HRTF (Head-Related Transfer Function – hàm truyền liên quan đến đầu).

Một cặp HRIR và chính nó nhìn theo tần số

- HRIR là cặp đáp ứng xung, sống trong miền thời gian.
- HRTF là cùng cặp ấy sau biến đổi Fourier, sống trong miền tần số.
- Nói về cài đặt dùng HRIR; nói về đặc tính nghe được dùng HRTF.



(a) Đáp ứng xung hai tai ở góc 40° (HRIR)



(b) Cùng cặp ấy nhìn theo tần số (HRTF)

Hình 10.21 — Một cặp HRIR và chính nó nhìn theo tần số

Tích chập với đáp ứng xung của từng tai

$$y_L[n] = \left(h_L^{\theta,\varphi} * x\right)[n]\,; \qquad y_R[n] = \left(h_R^{\theta,\varphi} * x\right)[n]$$

(10.13)

- Tích chập tín hiệu mono với đáp ứng xung tai trái để ra kênh trái.
- Tương tự với tai phải để ra kênh phải.
- θ và ϕ là phương vị và độ cao: đổi hướng nghĩa là đổi cặp đáp ứng xung.

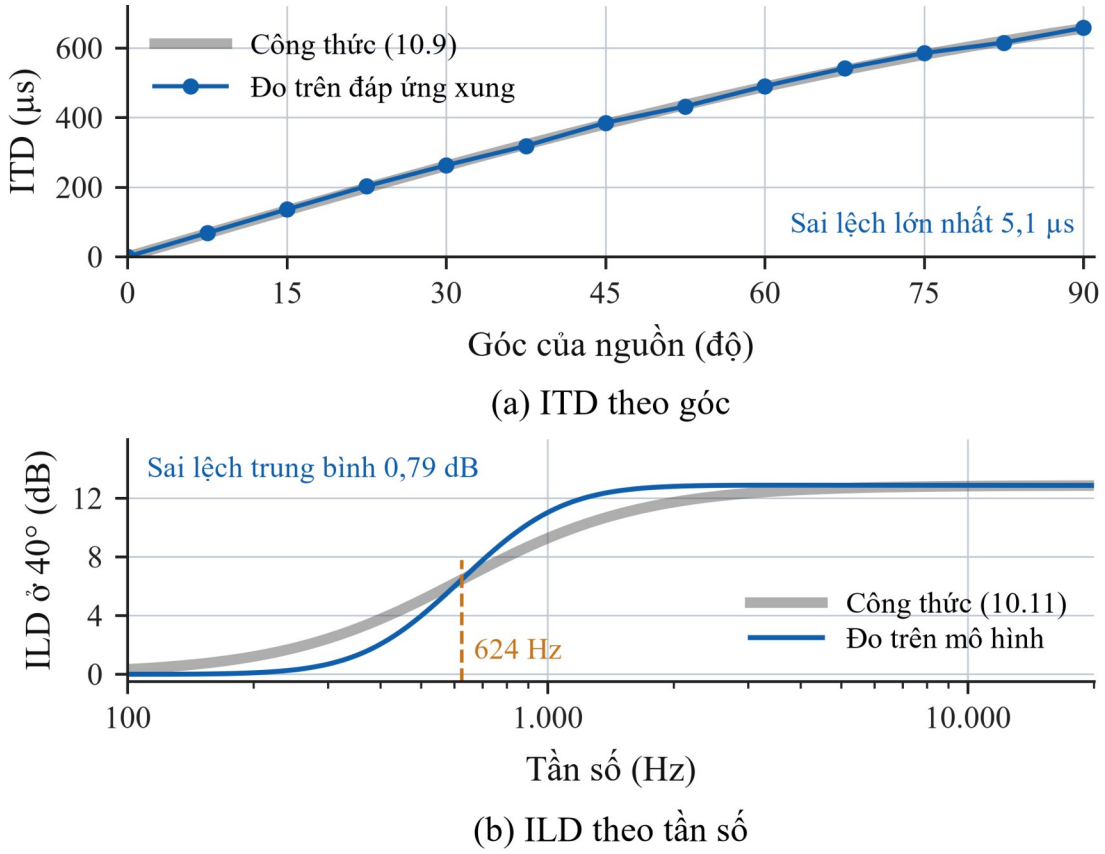
Đặt nguồn mono vào một hướng chỉ bằng ITD và ILD

```
def f_ka_bang_mot(ban_kinh=BAN_KINH_DAU, c=TOC_DO_AM):
    """Tần số mà ka bằng 1 — điểm uốn của công thức (10.11)."""
    return float(c / (2.0 * math.pi * float(ban_kinh)))

def binaural_don_gian(x, goc_do=30.0, sr=SR, ban_kinh=BAN_KINH_DAU):
    m = sang_mono(x)
    som = m                                        # tai gần
    muon = tre_phan_so(m, abs(itd_giay(goc_do, ban_kinh)) * sr,
                       'lagrange3')                # tai xa: trễ ITD
    do_bong = abs(math.sin(math.radians(float(goc_do))))
    b, a = biquad_he_so('ke_cao', f_ka_bang_mot(ban_kinh),
                        0.7071, -20.0 * do_bong, sr)
    muon = ap_biquad(muon, b, a)                   # tai xa: bóng đầu
    if goc_do >= 0:                                # nguồn bên phải
        return np.column_stack([muon, som])
    return np.column_stack([som, muon])
```

(rút gọn: bỏ phần ghi chú đầu hàm)

Đo lại ITD và ILD trên chính đáp ứng xung của mô hình



Hình 10.22 — Đo lại ITD và ILD trên chính đáp ứng xung của mô hình

Kết quả phép kiểm đặt rồi đo

Đại lượng	Sai lệch	Kết luận
ITD	Lớn nhất 4,8 µs trên cả dải góc	Khớp công thức (10.9), dưới một phần trăm mẫu
ILD	Trung bình 0,79 dB	Chỉ khớp đại khái: kệ bậc hai không phải đường cong bóng đầu

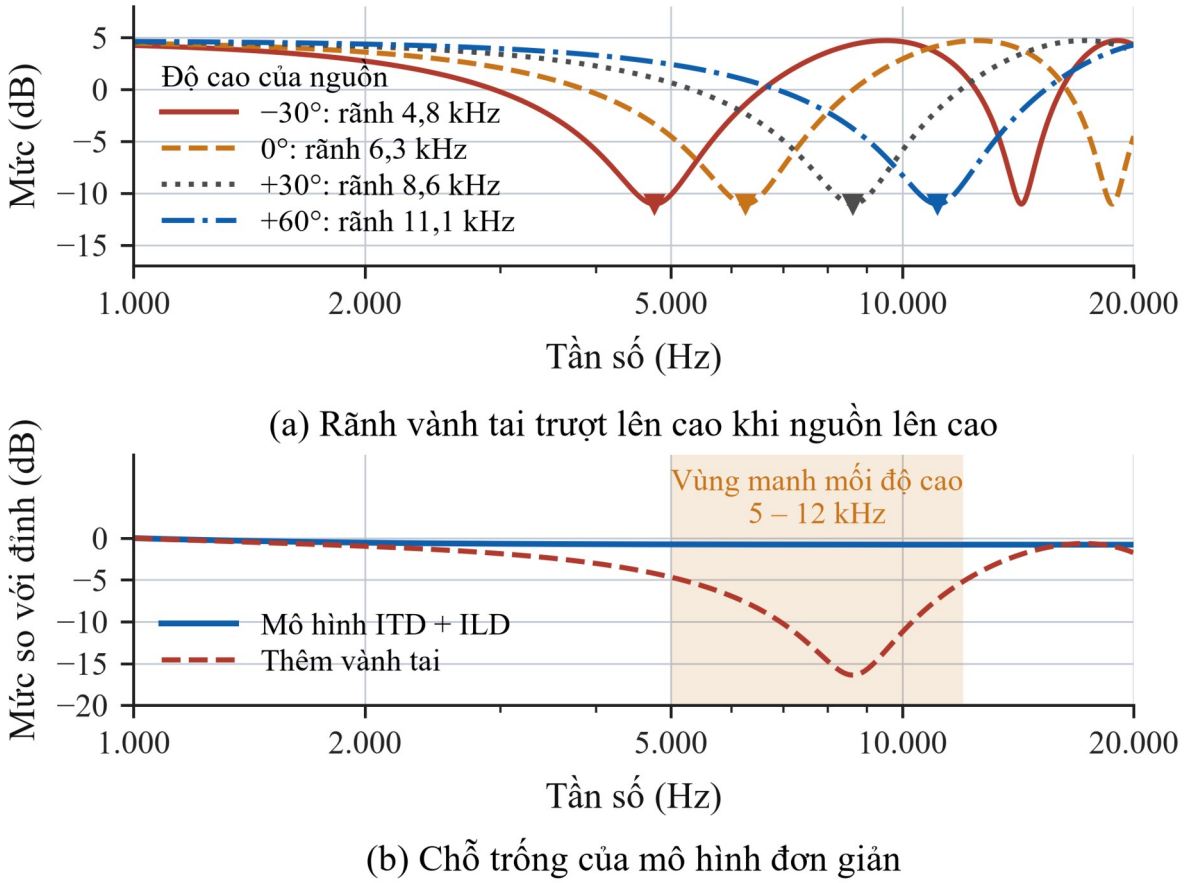
ITD đo bằng khoảng cách hai đỉnh, nội suy parabol; ILD đo bằng hiệu hai đường |H(f)| so với công thức.

Kiểm cài đặt không phải kiểm mô hình

Phép kiểm chỉ nói rằng mã làm đúng điều công thức yêu cầu, không nói công thức mô tả đúng tai người. Một mô hình sai vẫn có thể được cài đặt hoàn hảo.

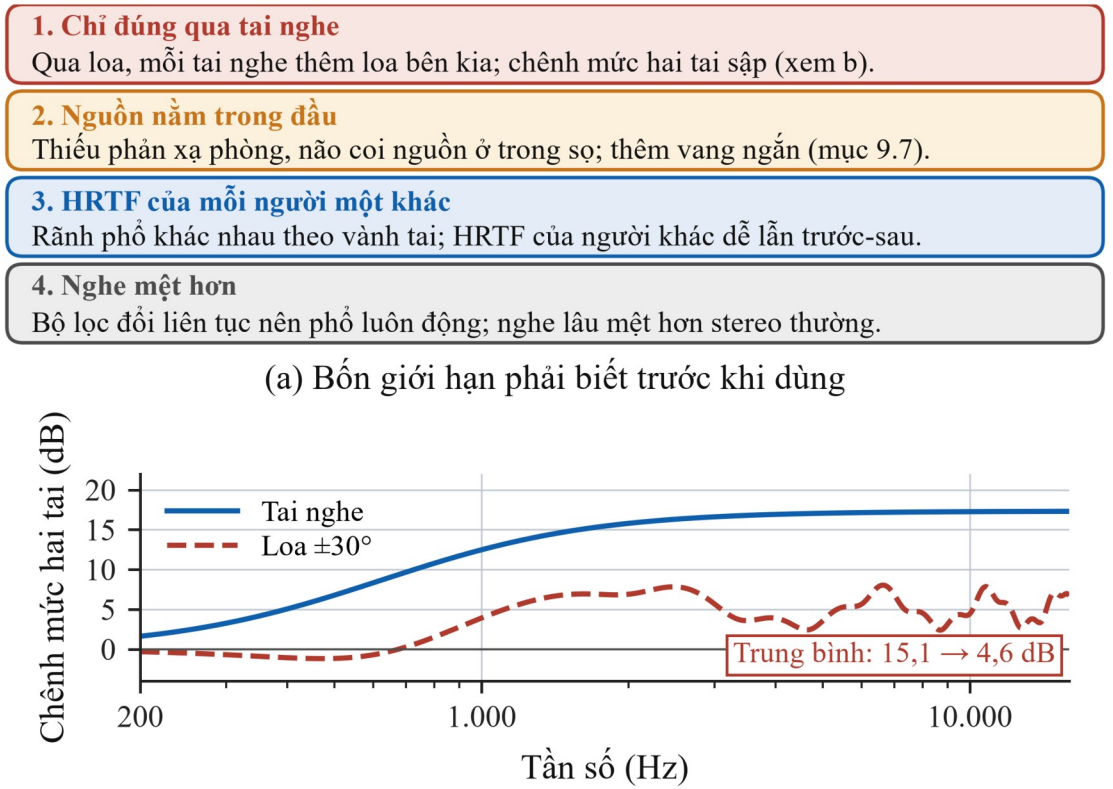
Rãnh phổ của vành tai trượt theo độ cao

- Mô hình đơn giản có ITD và ILD, thiếu phổ vành tai.
- Vì vậy nguồn nghe được trái phải nhưng không biết trên dưới.
- Âm tới thẳng vào ống tai, đồng thời một phần dội từ gờ vành tai.
- Hai đường cộng nhau là bộ lọc lược của mục 8.5.



Hình 10.23 — Rãnh phổ của vành tai trượt theo độ cao, và chỗ trống trong mô hình

Bốn giới hạn, và phép đo chứng minh giới hạn thứ nhất



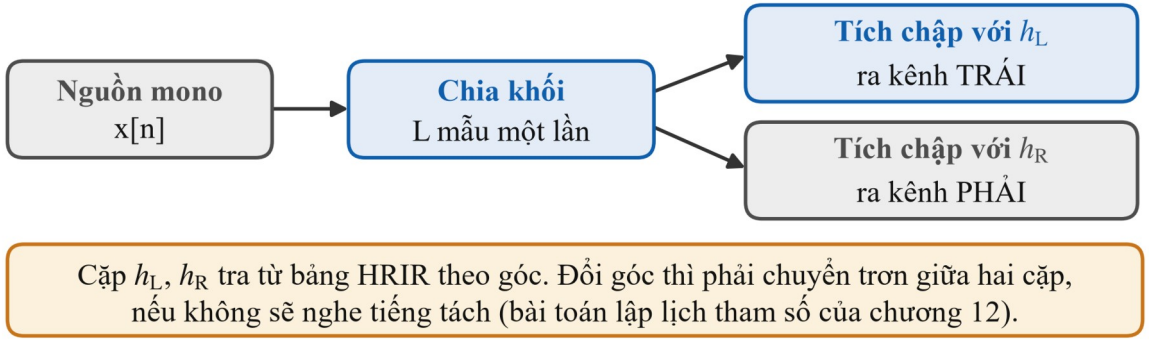
Hình 10.24 — Bốn giới hạn, và phép đo chứng minh giới hạn thứ nhất

Bốn giới hạn của âm thanh binaural

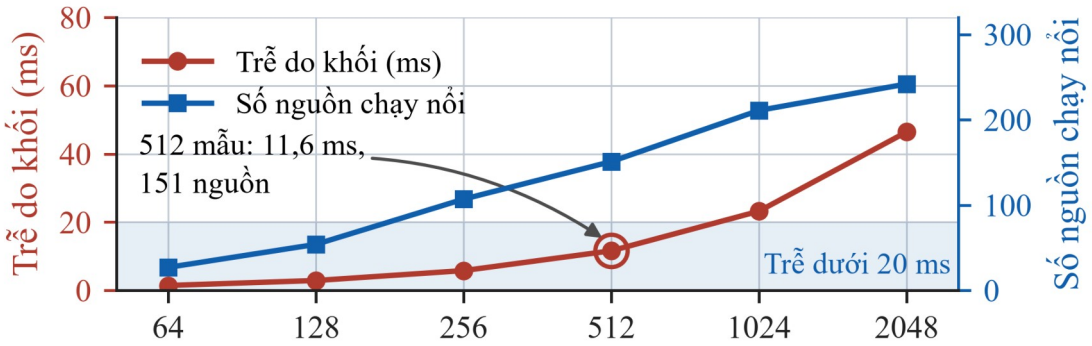
Giới hạn	Nội dung	Làm gì được
1 · Chỉ đúng qua tai nghe	Qua loa, ILD ở 60 độ sập từ 15,1 xuống 4,6 dB	Ghi rõ dùng tai nghe; hoặc khử xuyên âm
2 · Nguồn nằm trong đầu	Thiếu phản xạ phòng nên nguồn nghe trong sọ	Thêm vang ngắn mục 9.7
3 · HRTF mỗi người một khác	Dùng HRTF người khác thì hay lẫn trước/sau	Cho chọn bộ HRTF; thêm chuyển động đầu
4 · Nghe mệt hơn	Bộ lọc đổi liên tục, phổ luôn động	Chỉ dùng cho nguồn cần định vị

Đánh đổi giữa trễ và số nguồn

- Một nguồn là hai phép tích chập, mỗi phép với đáp ứng xung 256 mẫu.
- Dùng hàm tích chập theo khối của mục 9.8.
- Kích thước khối là tham số duy nhất phải chọn.



(a) Một nguồn đặt vào một hướng



(b) Khối càng lớn càng rẻ, nhưng càng chậm

Hình 10.25 — Sơ đồ một nguồn binaural và đánh đổi giữa trễ với số nguồn

Chọn kích thước khối

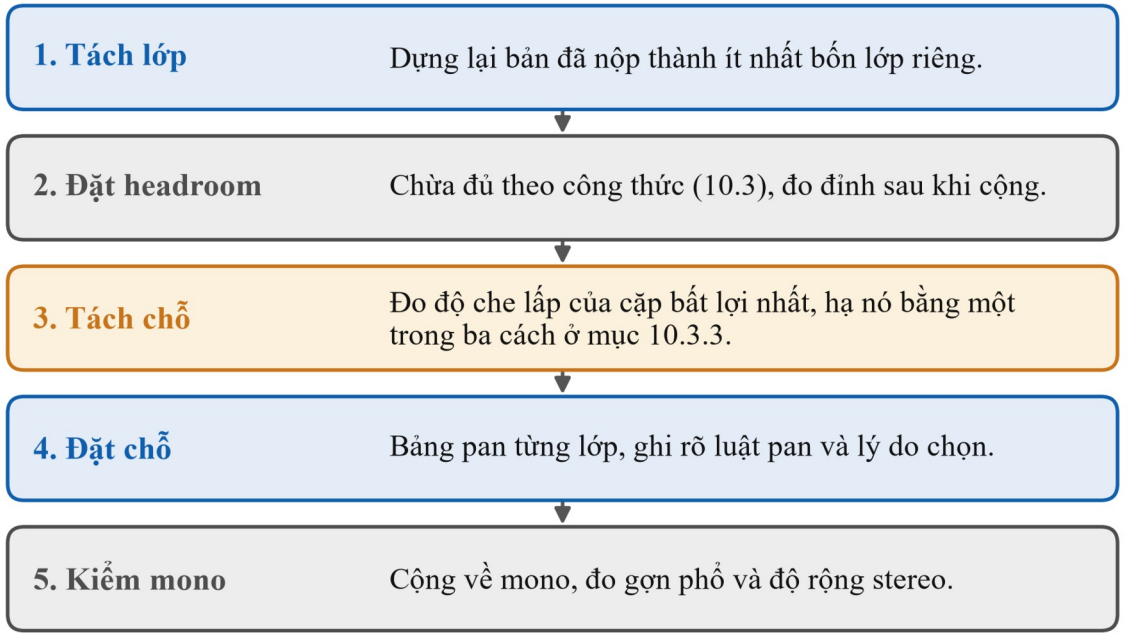
Khối	Trễ do khối	Số nguồn chạy nổi	Nhận xét
64 mẫu	1,5 ms	ít nhất	Trễ rất thấp; hợp nhạc cụ điện tử chơi trực tiếp
256 mẫu	5,8 ms	khá	Nhiều bộ máy trò chơi dừng ở đây
512 mẫu	11,6 ms	nhiều	Thường chọn: trễ dưới ngưỡng nhận biết
1024 mẫu trở lên	23 ms trở lên	gần như không tăng	Chỗ gãy: trễ gấp đôi mà số nguồn đứng yên

Số nguồn là số đo trên máy dựng giáo trình; máy khác ra con số khác, hình dạng đường cong không đổi.

MỤC 10.8

Mốc bài tập lớn — giai đoạn 3

- Nộp trước chương 13, làm tiếp trên bản đã nộp ở giai đoạn 2.
- Không cần tư liệu mới.
- Bản giai đoạn 2 chỉ có một lớp thì tách thành nhiều lớp, hoặc thu thêm ba đoạn ngắn.



Nộp: bốn tệp lớp rời + tệp trộn stereo + tệp trộn mono + bảng số đo + một trang giải thích.

Giai đoạn 3 không cần tư liệu mới. Nếu bản giai đoạn 2 chỉ có một lớp thì tách nó ra, hoặc thu thêm ba lớp ngắn.

Nội dung phải nộp ở giai đoạn 3

Hạng mục	Yêu cầu cụ thể
Bốn tệp lớp rời	Ít nhất bốn lớp, mỗi lớp một WAV, cùng tần số lấy mẫu và độ dài
Bảng headroom	N, lượng hạ mỗi lớp, đỉnh sau khi cộng; so với (10.3)
Bảng che lấp	Che lấp cặp bất lợi nhất trước và sau, kèm cách đã dùng
Bảng pan	Vị trí p, tên luật pan, lý do chọn
Tệp trộn stereo	Bản hoàn chỉnh, đỉnh không vượt −1,0 dBFS
Tệp trộn mono	Cùng bản ấy cộng về mono, kèm gợn phổ và mất mức
Một trang giải thích	Vì sao đặt mỗi lớp vào chỗ ấy; trọng số lớn nhất

Chín ô phải tự kiểm trước khi nộp

- Chín ô chấm bảy điểm và đều do máy kiểm được.
- Ba điểm còn lại nằm ở trang giải thích.
- Mỗi ô sai lệch quá ngưỡng mất khoảng tám phần mười điểm.

Hạng mục	Con số phải báo	Ngưỡng đạt
☐ Đỉnh của bản trộn stereo	**dBFS**	Không vượt −1,0
☐ Headroom đã chừa cho từng lớp	**dB, kèm N**	Khớp (10.3) ±1 dB
☐ Độ che lấp cặp lớp bất lợi nhất	**trước → sau**	Giảm ít nhất 0,10
☐ Luật pan đã dùng	**tên luật**	Nêu được lý do
☐ Vị trí pan từng lớp	**từ −1 đến 1**	Không quá 2 lớp ở giữa
☐ Độ rộng stereo của bản trộn	**0 đến 1**	Trong khoảng 0,10–0,60
☐ Gợn phổ khi cộng về mono	**dB**	Dưới 3,0
☐ Mất mức khi cộng về mono	**dB**	Dưới 3,0
☐ Độ to tích hợp của bản cuối	**LUFS**	Đúng đích GĐ2 ±0,5

Chín ô này chấm bảy điểm; ba điểm còn lại nằm ở trang giải thích: vì sao đặt mỗi lớp vào chỗ ấy.

Bản nào không kèm tệp mono thì mất trọn hai ô kiểm mono: kiểm mono là bắt buộc, không phải tuỳ chọn.

Hình 10.27 — Chín ô phải tự kiểm trước khi nộp

MỤC 10.9

Thực hành, bài tập và tổng kết

- Ba bài thực hành, mỗi bài chừng mười phút; phải nghe chứ không chỉ nhìn số.
- Bài thứ ba bắt buộc dùng tai nghe.
- Bảy bài tập có đáp số; điểm nằm ở cách làm và phần giải thích.

Headroom, pan với mono, Haas

- **Bài 1 · Headroom:** sáu tệp về −6 dBFS, cộng thẳng, hạ 12 dB, hạ theo (10.3); đo đỉnh và nghe cả ba.
- **Bài 2 · Pan và mono:** pan hai lớp xung đột bằng ba luật; đo mức ở giữa khi stereo và khi mono, khớp bảng 10.9.
- **Bài 3 · Haas:** làm rộng ở 0,5; 15; 40 ms; nghe stereo, nghe mono, đo gợn phổ; ghi lượng trễ tốt nhất cho cả hai.

Bảy bài tập: đề và đáp số (1–4)

#	Đề bài	Đáp số
1	Headroom cho 5 lớp độc lập; 5 lớp đồng pha	−6,99 dB; −13,98 dB
2	Cặp gain ba luật pan ở p = 0,5; mức qua loa và mức mono	Biên độ −2,04 và 0,00; công suất 0,00 và +2,32; thoả hiệp −1,03 và +1,15 dB
3	Rãnh đầu tiên khi mono của Haas 3 ms và 25 ms; lớp nào nguy hơn	167 Hz và 20,0 Hz
4	ITD cho nguồn 45 độ; ITD 300 µs ứng góc nào	380,7 µs; 34,7 độ

Bảy bài tập: đề và đáp số (5–7)

#	Đề bài	Đáp số
5	ILD nguồn 30 độ tại 300 Hz và 6 kHz; giải thích bằng ka	1,88 dB và 9,89 dB
6	Đầu trẻ em bán kính 7 cm: ITD lớn nhất, tần số giới hạn của ITD, tần số ka = 1	524,7 µs; 1.225 Hz; 780 Hz
7	Tích chập khối 512 với HRIR 256: so số phép nhân của cách thẳng và FFT	Thẳng 131.072; FFT 31.744; nhanh 4,13 lần

Bài 7: giải thích vì sao tỉ lệ nhỏ hơn hẳn con số 608 lần của mục 9.8.

Bảy câu không có đáp án in sẵn

- Vì sao RMS tám lớp độc lập tăng đúng công thức còn đỉnh thì ít hơn? Chỗ chênh dùng vào việc gì?
- Che lấp 0,68 giữa hai lớp nền và 0,54 có giọng hát: xử lý cặp nào trước, vì sao?
- Vì sao cắt giảm an toàn hơn nâng, dù cho cùng hiệu quả nghe?
- Chứng minh bằng lời: không luật pan nào phẳng được cả mức loa lẫn mức mono.
- Vì sao Haas trễ ngắn nguy hơn trễ dài khi về mono? Trả lời bằng khoảng cách giữa rãnh.
- ITD lớn nhất 656 µs, tai phân biệt 10 µs: ý nghĩa với cách viết phép trễ?
- Mô hình binaural đơn giản thiếu gì để phân biệt trên với dưới, và vì sao?

Sáu câu phải trả lời được sau chương 10

Chừa headroom bao nhiêu cho N lớp?
Mười lần log của N nếu các lớp độc lập. Tám lớp đo được +8,9 dB RMS, đúng như (10.3).

Vì sao đỉnh lại tăng ÍT hơn công thức?
Tám lớp chỉ đẩy đỉnh lên 7,2 dB. Phần chênh ấy là chỗ bộ giới hạn của chương 8 làm việc.

Hai lớp xung đột thì làm gì?
Đo độ che lấp rồi hạ bằng một trong ba cách. Cặp bất lợi nhất của bài giảm từ 0,68 xuống 0,41.

Luật pan nào đúng?
Không luật nào đúng cả hai đầu. Giữ biên độ thì hụt 3,01 dB công suất ở giữa; giữ công suất thì tăng thêm 3,01 dB khi về mono.

Haas nguy ở chỗ nào?
Ở trễ NGẮN. Cộng về mono, bản 0,5 ms gợn 7,5 dB với rãnh sâu −40,6 dB; bản 40 ms chỉ gợn 1,0 dB.

Vì sao binaural phải nghe bằng tai nghe?
Qua loa, tiếng xuyên chéo kéo chênh mức hai tai từ 15,1 xuống 4,6 dB, gần hết thông tin hướng.

Chương 11 mở mô-đun M7: đưa các khối này vào thời gian thực (Web Audio).

Hình 10.28 — Sáu câu phải trả lời được sau chương 10

Những hàm mới của chương và chỗ dùng lại

Hàm	Làm gì	Dùng lại ở
cong_lop	Cộng nhiều lớp, mỗi lớp một gain	Chương 12
headroom_can_db	Lượng hạ mỗi lớp cho N lớp	Bài tập lớn giai đoạn 3
he_so_tuong_quan	Tương quan Pearson giữa hai lớp	Dò đảo cực; độ rộng stereo
che_lap_pho	Độ che lấp phổ, trong [0, 1]	Chương 13
luat_pan · pan	Ba luật pan và phép áp	Chương 11
do_rong_stereo	Độ rộng stereo bằng tương quan	Chương 14
gon_pho_khi_ve_mono_db	Gợn phổ, rãnh sâu nhất khi mono	Giai đoạn 3 và 4

Rút gọn từ Bảng 10.23 (7 trong 10 dòng; lược haas, itd_giay · ild_db, binaural_don_gian).

Chương 8, 9, 10 dựng xong bộ công cụ xử lý

Khoảng động, tần số, thời gian và không gian đều đã có, nhưng vẫn chạy theo lô. Chương 11 mở mô-đun M7: cùng các khối ấy, chạy thời gian thực trong trình duyệt, phản ứng theo sự kiện.

Những điều cần nhớ

1. Trộn là cộng có trọng số; headroom cho N lớp độc lập là mười lần loga của N, đồng pha gấp đôi.
2. Hệ số tương quan dự đoán mức tăng khi cộng và phát hiện lớp bị đảo cực.
3. Độ che lấp phổ xếp thứ tự việc; vai trò của lớp quyết định việc nào làm trước.
4. Ba luật pan: không luật nào phẳng cả khi nghe loa lẫn khi cộng mono.
5. Haas trễ ngắn nguy hơn khi về mono vì rãnh lọc lược thưa và rơi giữa vùng nghe.
6. ITD cho dải trầm, ILD cho dải cao; binaural cần HRTF và có bốn giới hạn phải biết.