MỤC LỤC

Chương 10. Trộn nhiều lớp và âm thanh không gian 5
10.1 Vị trí của chương trong cuốn sách 6
10.2 Trộn nhiều lớp và headroom 8
10.3 Che lấp phổ và ba cách tách lớp 13
10.4 Panning — đặt lớp vào chỗ giữa hai loa 18
10.5 Hiệu ứng Haas — làm rộng bằng thời gian 23
10.6 ITD và ILD — tai người định vị bằng gì 27
10.7 Binaural và HRTF 31
10.8 Dự án thực hành — giai đoạn 3 37
10.9 Thực hành, bài tập và tổng kết 39

DANH MỤC HÌNH VẼ

Hình 10.1 Bảy mục kỹ thuật của chương 10.....6
Hình 10.2 Từ nhiều lớp rời tới một bản trộn nghe được.....6
Hình 10.3 Bốn kiểu quan hệ giữa hai lớp và mức tăng tương ứng.....9
Hình 10.4 Hai công thức là hai đầu mút, và chỗ thực tế nằm giữa.....11
Hình 10.5 Bảng tra headroom theo số lớp.....11
Hình 10.6 Sáu lớp: cộng thẳng, hạ theo cảm tính, và hạ theo công thức.....12
Hình 10.7 Phổ của sáu lớp vẽ chồng lên nhau.....14
Hình 10.8 Ma trận che lấp của sáu lớp và sáu cặp xung đột nhiều nhất.....15
Hình 10.9 Tách bằng tần số, bằng không gian và bằng chiều sâu.....16
Hình 10.10 Xếp thứ tự việc phải làm theo độ che lấp và theo tầm quan trọng.....17
Hình 10.11 Ba luật pan nhìn theo đường gain.....19
Hình 10.12 Cái giá của từng luật ở hai đầu: nghe loa và cộng mono.....20
Hình 10.13 Pan cặp xung đột ra hai phía và đo lại độ tách.....21
Hình 10.14 Sáu lớp trên bản đồ pan và bốn con số đo lại.....22
Hình 10.15 Hai kênh cùng mức, chỉ khác thời điểm — và bảng năm mốc trễ.....24
Hình 10.16 Gợn phổ khi cộng về mono, theo lượng trễ.....25
Hình 10.17 Bốn cách làm rộng một lớp, được gì và mất gì.....26
Hình 10.18 ITD theo góc, và độ phân giải của tai.....27
Hình 10.19 ILD theo tần số và theo góc — thuyết song công.....29
Hình 10.20 Nón lẫn lộn, và độ nhạy hướng theo góc.....30
Hình 10.21 Một cặp HRIR và chính nó nhìn theo tần số.....31
Hình 10.22 Đo lại ITD và ILD trên chính đáp ứng xung của mô hình.....32
Hình 10.23 Rãnh phổ của vành tai trượt theo độ cao, và chỗ trống trong mô hình.....33
Hình 10.24 Bốn giới hạn, và phép đo chứng minh giới hạn thứ nhất.....34
Hình 10.25 Sơ đồ một nguồn binaural và đánh đổi giữa trễ với số nguồn.....36
Hình 10.26 Năm bước của giai đoạn 3.....37
Hình 10.27 Chín ô phải tự kiểm trước khi coi là xong.....38
Hình 10.28 Sáu câu phải trả lời được sau chương 10.....41

DANH MỤC BẢNG BIỂU

Bảng 10.1 Bảy mục kỹ thuật của chương 10 dựng trên chương nào....7
Bảng 10.2 Ba thứ dễ gộp nhầm vào chương 10....8
Bảng 10.3 Bốn trường hợp và mức tăng khi cộng hai lớp cùng mức....10
Bảng 10.4 Headroom cần chừa cho từng lớp....12
Bảng 10.5 Ba cách xử lý cùng sáu lớp ấy....13
Bảng 10.6 Ba cặp xung đột nhiều nhất trong sáu lớp của bài....15
Bảng 10.7 Ba cách tách và cái giá của từng cách....16
Bảng 10.8 Danh sách việc rút ra từ hai bước xếp thứ tự....18
Bảng 10.9 Ba luật pan và cái giá của mỗi luật....20
Bảng 10.10 Bốn quy ước đặt chỗ và lý do của từng quy ước....22
Bảng 10.11 Bốn con số đo trên bản trộn trước và sau khi đặt chỗ....23
Bảng 10.12 Trễ bao nhiêu thì nghe ra gì....24
Bảng 10.13 Trễ và cái giá tương ứng khi cộng về mono....25
Bảng 10.14 Ba manh mối định vị và vùng tần số của từng manh mối....27
Bảng 10.15 Bốn cách giải quyết nón lẫn lộn....30
Bảng 10.16 Kết quả phép kiểm đặt rồi đo....33
Bảng 10.17 Rãnh vành tai trượt theo độ cao — mô hình phản xạ đơn giản hoá....34
Bảng 10.18 Bốn giới hạn của âm thanh binaural....35
Bảng 10.19 Chọn kích thước khối — đánh đổi giữa hai thứ....36
Bảng 10.20 Kết quả cần có ở giai đoạn 3....38
Bảng 10.21 Bảy bài tập của chương 10....40
Bảng 10.22 Những hàm mới của chương 10 và chỗ dùng lại....42

CHƯƠNG 10
TRỘN NHIỀU LỚP VÀ ÂM THANH KHÔNG GIAN

Mục tiêu của chương

Sau chương này, người đọc phải làm được những việc sau.

- Tính được lượng headroom cần chừa cho N lớp, và giải thích được vì sao có hai công thức chứ không phải một.
- Dùng hệ số tương quan để dự đoán mức tăng khi cộng, và để phát hiện một lớp bị đảo cực.
- *đo* được độ che lấp phổ giữa hai lớp và xếp thứ tự việc phải làm theo con số ấy.
- Chọn được một trong ba cách tách lớp và nói rõ cái giá của cách đã chọn.
- Cài đặt được cả ba luật pan, và chứng minh được rằng không luật nào đúng ở cả hai đầu.
- Dùng được hiệu ứng Haas, và *đo* được cái giá của nó khi bản trộn bị cộng về mono.
- Tính được ITD và ILD cho một hướng bất kỳ, và giải thích được thuyết song công bằng đại lượng ka.
- Dựng được một mô hình binaural đơn giản từ hai khối đã có của chương 9, rồi *đo* lại xem cài đặt có khớp công thức không.
- Nêu được bốn giới hạn của binaural và biết giới hạn nào đo được bằng số.
- Chọn được kích thước khối cho một bộ máy binaural thời gian thực, cân giữa độ trễ và số nguồn.

Hai câu hỏi, và chúng dẫn tới hai nửa của chương

Nửa đầu — mục 10.2 tới 10.4 — hỏi: nhiều lớp cộng lại thì to lên bao nhiêu, và làm sao để chúng không lấn nhau. Đây là bài toán của người trộn nhạc.

Nửa sau — mục 10.5 tới 10.7 — hỏi: tai người dựa vào đâu để biết một tiếng đến từ hướng nào. Đây là bài toán của người làm trò chơi và thực tế ảo.

Hai nửa gặp nhau ở một chỗ: mọi kỹ thuật của nửa đầu đều hiệu quả vì chúng mô phỏng đúng những manh mối của nửa sau.

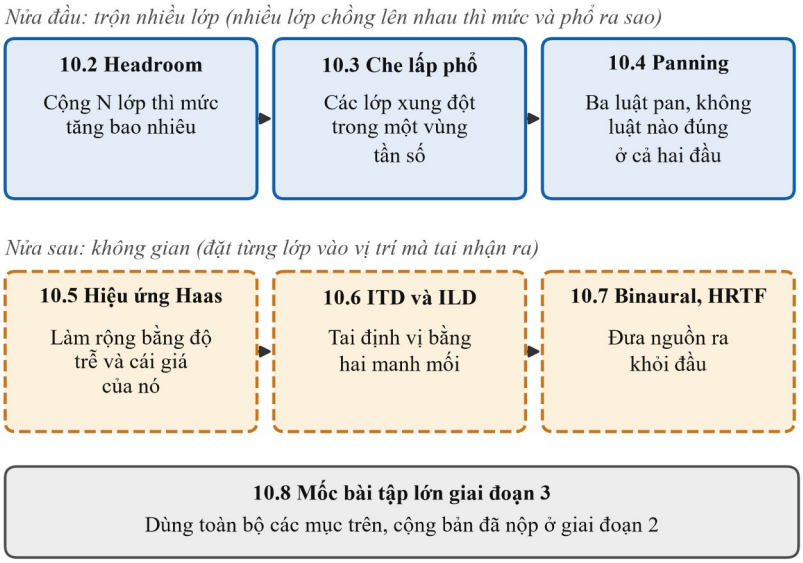
Ba mục có phép kiểm bằng số nên tự chạy lại: headroom ở mục 10.2.4, tách cặp che lấp ở mục 10.3.3, và cài đặt binaural ở mục 10.7.2.

10.1 Vị trí của chương trong cuốn sách

10.1.1 Chín chương vừa qua chỉ làm việc với một tín hiệu

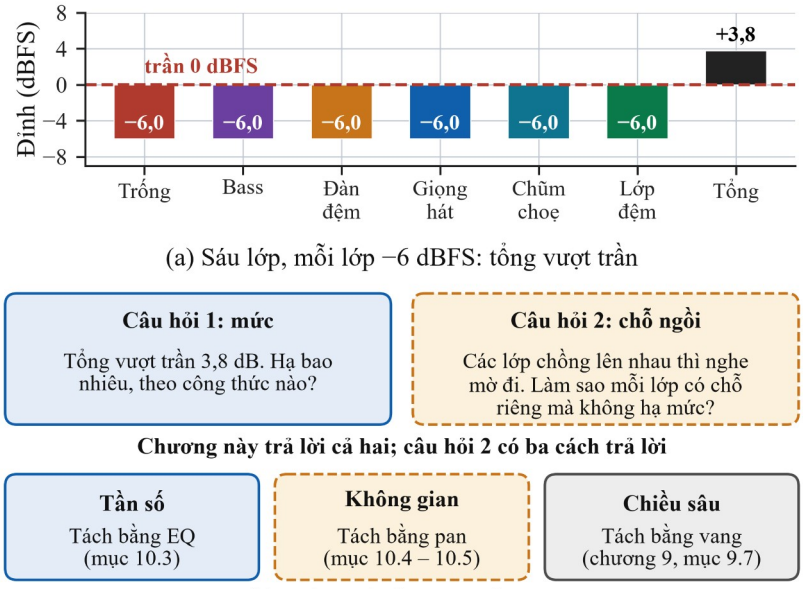
Từ chương 5 tới chương 9, mọi thứ đã học đều nhận vào một tín hiệu và trả ra một tín hiệu. Chuẩn hoá, khử nhiễu, nén, giới hạn, trễ, lọc, vang — không phép nào trong số ấy hỏi xem còn có tín hiệu nào khác đang phát cùng lúc hay không.

Một sản phẩm thật thì không như vậy. Nó có trống, có bass, có giọng hát, có nền — nhiều lớp phát cùng lúc, và người nghe có hai tai chứ không phải một. Chương này bổ sung đúng hai điều ấy: cộng nhiều lớp thì phải chừa chỗ bao nhiêu, và đặt mỗi lớp vào chỗ nào trong không gian giữa hai tai.



Hình 10.1 Bảy mục kỹ thuật của chương 10

10.1.2 Bài toán của chương



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

(b) Hai câu hỏi của cả chương

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

Bài toán của chương phát biểu gọn như sau. Cho N lớp âm thanh, mỗi lớp đã xử lý xong bằng công cụ của các chương trước. Hãy chọn gain và vị trí cho từng lớp sao cho bản trộn không cắt đỉnh, các lớp không lấn nhau, và bản trộn ấy vẫn nghe được khi ai đó cộng nó về mono.

Ba yêu cầu ấy mâu thuẫn nhau từng đôi một, và phần lớn nội dung chương này là học cách đo mức mâu thuẫn để chọn có căn cứ chứ không chọn theo cảm tính.

10.1.3 Bảy mục và thứ mà mỗi mục dựng trên

Bảng 10.1 Bảy mục kỹ thuật của chương 10 dựng trên chương nào

Mục	Nội dung	Dựng trên
10.2	Cộng nhiều lớp và headroom	Phép cộng và gain theo decibel của chương 5; bộ giới hạn của mục 8.4
10.3	Che lấp phổ và ba cách tách lớp	Phổ của chương 4; bộ EQ của mục 9.4; bộ vang của mục 9.7
10.4	Panning và ba luật pan	Hoàn toàn mới
10.5	Hiệu ứng Haas	Phép trễ của mục 6.4 và bộ lọc lược của mục 8.5
10.6	ITD, ILD và thuyết song công	Trễ phân số của mục 9.2; đáp ứng biên độ của mục 9.3
10.7	Binaural và HRTF	Tích chập theo khối của mục 9.8
10.8	Dự án thực hành giai đoạn 3	Toàn bộ bảy mục trên, cộng bản đã làm ở giai đoạn 2

Một câu tóm cả chương

Chương 8 và chương 9 hỏi: tín hiệu *này* nghe ra sao.

Chương 10 hỏi hai câu khác: nhiều tín hiệu cộng lại thì to lên bao nhiêu, và tai người dựa vào đâu để biết một tiếng đến từ hướng nào.

Câu thứ nhất trả lời bằng một công thức duy nhất — công thức (10.3). Câu thứ hai trả lời bằng ba manh mối, và cả ba đều đo được bằng số.

10.1.4 Những gì chương này không làm

Bảng 10.2 Ba thứ dễ gộp nhầm vào chương 10

Thứ dễ gộp nhầm	Vì sao không thuộc chương này	Thuộc chương nào
Trộn động — mức các lớp thay đổi theo sự kiện trong lúc chạy	Chương này chọn *một* bộ gain và giữ nguyên. Đổi gain trong lúc chạy mà không nghe tiếng tách là bài toán lập lịch tham số	Chương 12
Âm thanh vòm nhiều loa — 5.1, 7.1, Atmos	Chúng cần thêm luật phân phối công suất cho nhiều hơn hai loa. Chương này chỉ làm hai kênh, vì hai kênh đã đủ để nêu mọi mâu thuẫn	Ngoài phạm vi sách
Chuẩn hoá độ to của bản trộn cuối cùng	Đó là mục 7.3, làm *sau* khi trộn xong. Chương này dừng ở chỗ có một bản trộn không cắt đỉnh	Chương 7

10.2 Trộn nhiều lớp và headroom

10.2.1 Phép trộn chỉ là một phép cộng có trọng số

Về mặt toán, trộn không có gì mới so với chương 5. Mỗi lớp nhân với một gain rồi cộng lại, đúng một dòng.

$$y[n] = \sum_{k=1}^{N} g_k x_k[n] \tag{10.1}$$

Trong đó x_k là lớp thứ k, g_k là gain của nó tính theo hệ số nhân (không phải theo decibel), và N là số lớp. Toàn bộ khó khăn không nằm ở phép cộng mà nằm ở chỗ chọn g_k: cộng N tín hiệu đầy biên độ thì kết quả vượt trần, và vượt trần là cắt đỉnh chứ không phải to lên.

Liệt kê 10.1 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
```

Ba lỗi cài đặt hay gặp ở hàm này

Cộng gain theo *decibel* thay vì nhân hệ số. Hai lớp cùng −6 dB không cho −12 dB; decibel chỉ cộng được khi nó là số mũ, còn ở đây phải đổi sang hệ số nhân trước.

Không đệm cho bằng độ dài. Lớp ngắn hơn bị numpy báo lỗi hình dạng, hoặc nghiêm trọng hơn là bị cắt cụt âm thầm nếu ai đó dùng phép cắt.

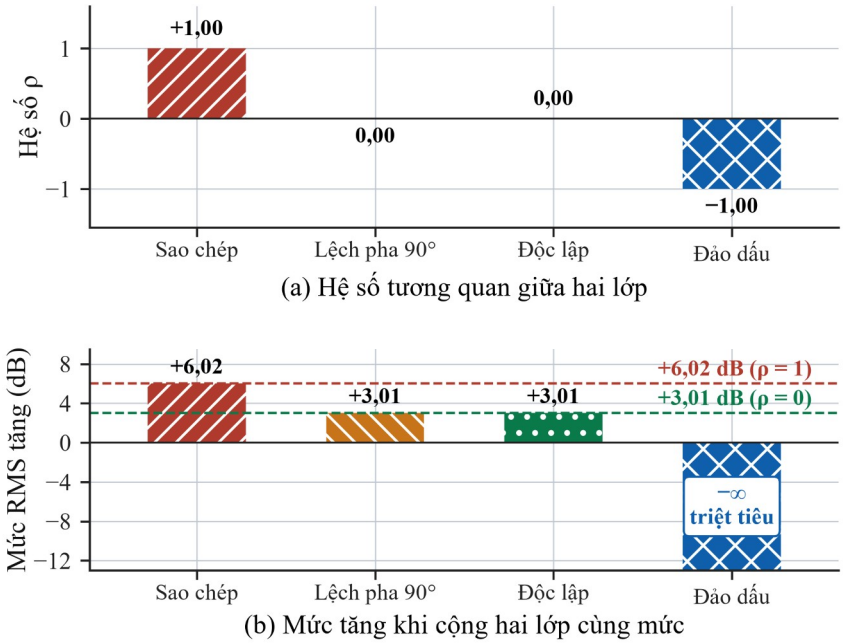
Trộn lớp mono với lớp stereo mà không lên kênh trước. Kết quả chạy được nhưng lớp mono chỉ vào một kênh.

10.2.2 Hai lớp cộng lại to lên bao nhiêu — câu trả lời phụ thuộc một con số

Câu hỏi tưởng đơn giản: cộng hai tín hiệu cùng mức thì mức tăng bao nhiêu decibel. Câu trả lời không phải một con số mà là một khoảng, và con số quyết định nằm ở khoảng ấy là hệ số tương quan giữa hai tín hiệu.

$$\rho = \frac{\sum_{n}(a[n]-\bar{a})(b[n]-\bar{b})}{\sqrt{\sum_{n}(a[n]-\bar{a})^{2}\sum_{n}(b[n]-\bar{b})^{2}}} \tag{10.2}$$

Đây là hệ số tương quan Pearson, chạy từ −1 tới +1. Bằng 0 nghĩa là hai lớp không liên quan gì tới nhau; bằng +1 nghĩa là chúng hoàn toàn giống nhau về hình dạng; bằng −1 nghĩa là lớp này là bản đảo dấu của lớp kia.



(a) Hệ số tương quan giữa hai lớp

(b) Mức tăng khi cộng hai lớp cùng mức

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

Bảng 10.3 Bốn trường hợp và mức tăng khi cộng hai lớp cùng mức

Quan hệ giữa hai lớp	ρ	Mức RMS tăng	Gặp ở đâu
Cùng một tín hiệu, nhân đôi	+1,00	+6,02 dB	Nhân bản một track để "dày" hơn — cách làm cạn headroom nhanh nhất
Cùng nguồn nhưng lệch pha 90 độ	0,00	+3,01 dB	Hai micro đặt lệch nhau, hoặc một lớp qua bộ thông toàn phần
Hai lớp độc lập	≈ 0,00	+3,01 dB	Trường hợp thường gặp nhất: hai nhạc cụ khác nhau
Bản gốc và bản đảo dấu	−1,00	triệt tiêu hoàn toàn	Đấu nhầm dây cân bằng, hoặc một plugin đảo cực mà không ai để ý

Cột cuối là cột nguy hiểm nhất

Khi ρ bằng −1, hai lớp không nhỏ đi mà *biến mất*. Trên hình, cột ấy chạy xuống âm vô cùng.

Trong studio, đây là lỗi hay gặp: một dây cân bằng đấu nhầm chân 2 với chân 3, hoặc một plugin có nút đảo cực bị bật. 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 nào cho ρ âm sâu là cặp có vấn đề.

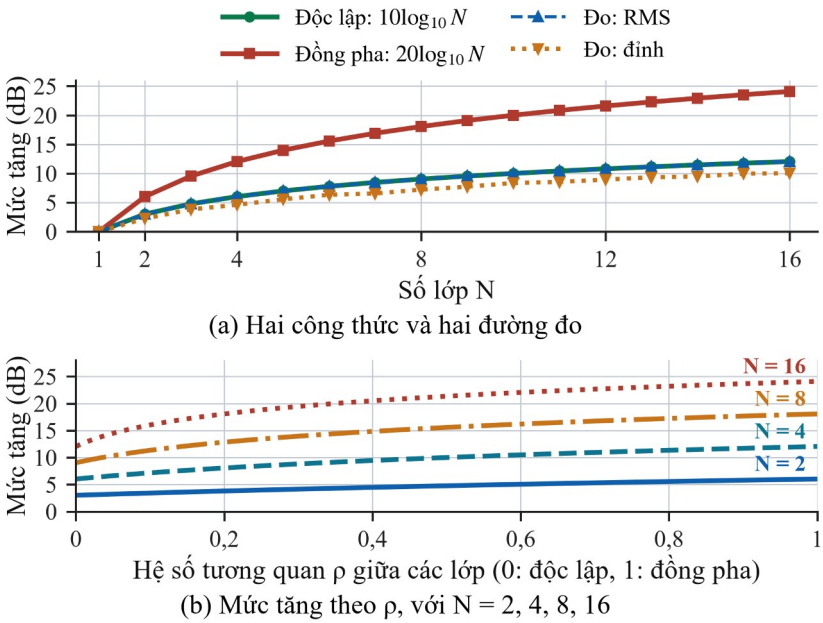
10.2.3 Công thức headroom

Từ hệ số tương quan ra thẳng công thức cần dùng. Với N lớp cùng mức, mức tổng tăng theo hai đầu mút sau.

$$\Delta L = 10\log_{10} N \ (\rho = 0); \qquad \Delta L = 20\log_{10} N \ (\rho = 1) \tag{10.3}$$

Đầu mút thứ nhất — các lớp độc lập — là trường hợp thực tế của gần như mọi bản trộn nhạc: công suất cộng lại, nên mức tăng theo mười lần loga. Đầu mút thứ hai — các lớp đồng pha — là trường hợp bất lợi nhất: biên độ cộng lại, nên mức tăng gấp đôi con số kia.

Muốn cộng N lớp mà không vượt trần thì hạ mỗi lớp đi đúng lượng ấy trước khi cộng. Đó chính là headroom cần chừa.



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

Hình (a) có bốn đường. Hai đường liền nét là hai công thức. Hai đường đứt nét là kết quả *đo* trên nhiễu độc lập, trung bình hai mươi tư lần cho mỗi số lớp.

Điều đáng chú ý nhất ở hình này

Đường RMS đo được bám sát đường lý thuyết độc lập — công thức đúng.

Nhưng đường *đỉnh* thì thấp hơn hẳn. Với tám lớp nhiễu độc lập, RMS tăng 8,9 decibel còn đỉnh chỉ tăng 7,2 decibel.

Lý do: đỉnh của các lớp hiếm khi rơi đúng vào cùng một mẫu. Vì thế chừa headroom theo công thức là *an toàn* nhưng hơi thừa — thừa chừng một tới hai decibel.

Chỗ thừa ấy không phí. Nó chính là dư địa cho bộ giới hạn của mục 8.4 làm việc mà không phải hạ thêm gì.

Headroom cần chừa cho mỗi lớp

Số lớp N	Độc lập (ρ = 0)	Đồng pha (ρ = 1)	Chênh lệch
1	0,0 dB	0,0 dB	0,0 dB
2	−3,0 dB	−6,0 dB	3,0 dB
4	−6,0 dB	−12,0 dB	6,0 dB
8	−9,0 dB	−18,1 dB	9,0 dB
12	−10,8 dB	−21,6 dB	10,8 dB
16	−12,0 dB	−24,1 dB	12,0 dB

Gấp đôi số lớp chỉ thêm 3 dB (cột độc lập).

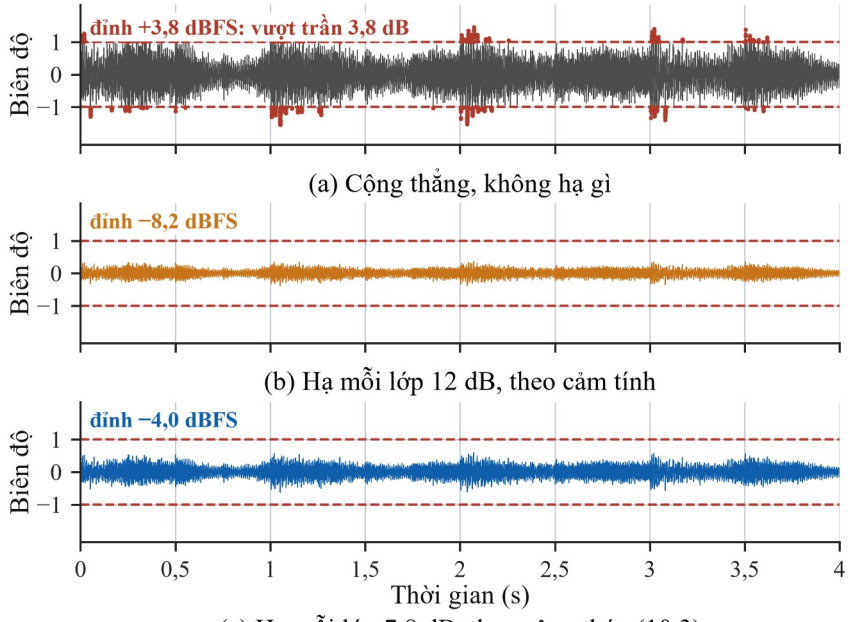
Hình 10.5 Bảng tra headroom theo số lớp

Bảng 10.4 Headroom cần chừa cho từng lớp

Số lớp N	Nếu độc lập	Nếu đồng pha	Ghi chú thực dụng
1	0,0 dB	0,0 dB	Không phải chừa gì
2	−3,0 dB	−6,0 dB	Hai lớp là chỗ chênh lệch còn nhỏ
4	−6,0 dB	−12,0 dB	Bốn lớp là mức tối thiểu của dự án thực hành giai đoạn 3
8	−9,0 dB	−18,1 dB	Bản trộn nhỏ điển hình
12	−10,8 dB	−21,6 dB	Chú ý mức tăng chậm dần: gấp đôi số lớp chỉ thêm 3 decibel
16	−12,0 dB	−24,1 dB	Vượt mốc này thì nên gom nhóm trước rồi trộn nhóm

Cột cuối chứa một nhận xét quan trọng về hình dạng của công thức: mười lần loga tăng rất chậm. Từ tám lớp lên mười sáu lớp — gấp đôi — chỉ tốn thêm ba decibel. Nghĩa là thêm vài lớp nữa vào một bản trộn đã đông không tốn mấy headroom; thứ đắt là bốn lớp đầu tiên.

10.2.4 Thử trên một bản trộn thật



(a) Cộng thẳng, không hạ gì

(b) Hạ mỗi lớp 12 dB, theo cảm tính

(c) Hạ mỗi lớp 7,8 dB, theo công thức (10.3)

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

Sáu lớp trong hình là sáu lớp sẽ dùng suốt chương: trống, bass, đàn đệm, giọng hát, chũm choẹ và lớp đệm. Mỗi lớp đã chuẩn hoá về −6 dBFS, đúng như một người mới trộn hay làm.

Bảng 10.5 Ba cách xử lý cùng sáu lớp ấy

Cách làm	Kết quả	Nhận xét
Cộng thẳng, không hạ gì	Vượt trần, cắt đỉnh nhiều đoạn	Đây là điều xảy ra khi ai đó kéo sáu track vào phần mềm rồi phát lại
Hạ theo cảm tính mỗi lớp 12 dB	Không cắt, nhưng bản trộn nhỏ hẳn	An toàn nhưng phí gần bốn decibel dải động; sau đó phải nâng lại, và nâng lại thì nâng cả nhiễu nền
Hạ đúng 7,8 dB theo công thức (10.3)	Không cắt, còn dư địa vừa phải	Đúng lượng cần: sáu lớp độc lập cho mười lần loga của sáu, bằng 7,78 decibel

Một phép kiểm nên tự chạy

Lấy sáu tệp bất kỳ, chuẩn hoá cả sáu về −6 dBFS, rồi làm ba lần: cộng thẳng, hạ 12 decibel, hạ 7,8 decibel.

Đo đỉnh của cả ba kết quả. Chỉ mất vài dòng, nhưng nó biến công thức (10.3) từ một dòng phải nhớ thành một thứ đã tự tay kiểm.

Nếu đỉnh của bản hạ 7,8 decibel vẫn vượt trần thì các lớp của bạn không độc lập — nhiều khả năng chúng cùng gốc, hoặc có hai lớp là bản sao của nhau.

10.2.5 Ba quy tắc rút ra

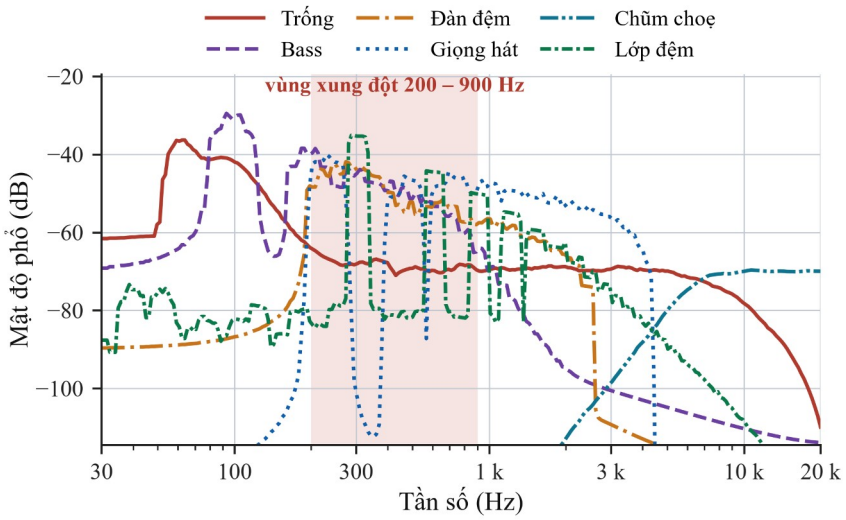
- **Chừa headroom trước, đừng sửa sau.** Hạ từng lớp trước khi cộng thì dải động của từng lớp còn nguyên. Cộng trước rồi hạ cả bản trộn thì phần đã cắt đỉnh không lấy lại được.
- **Đo hệ số tương quan khi con số không khớp.** Nếu đỉnh cao hơn công thức dự đoán nhiều thì các lớp đang tương quan với nhau. Đó là thông tin có ích, không phải phiền toái.
- **Đừng chừa thừa để cho chắc.** Mỗi decibel chừa thừa là một decibel phải nâng lại ở cuối, và lúc nâng lại thì nhiễu nền của mục 7.4 nâng theo.

10.3 Che lấp phổ và ba cách tách lớp

10.3.1 Vấn đề: đủ headroom rồi mà vẫn nghe đục

Giả sử đã chừa headroom đúng theo mục trước. Bản trộn không cắt đỉnh, mỗi lớp nghe riêng đều rõ, nhưng cộng lại thì nghe đục và không phân biệt được lớp nào với lớp nào. Đây là than phiền phổ biến nhất về một bản trộn, và nó không liên quan gì tới mức.

Nguyên nhân là các lớp chiếm cùng một vùng tần số. Tai người phân giải theo dải chứ không theo từng tần số riêng, nên khi hai lớp cùng đổ năng lượng vào một dải thì lớp mạnh hơn che lớp yếu hơn.



Hình 10.7 Phổ của sáu lớp vẽ chồng lên nhau

Hình trên là thứ nên mở ra trước khi chỉnh bất cứ điều gì. Đọc nó thì thấy ngay ba nhóm: trống và bass nằm dưới hai trăm hertz; chũm choẹ nằm trên sáu kilohertz; ba lớp còn lại — đàn đệm, giọng hát và lớp đệm — chen chúc trong vùng hai trăm tới chín trăm hertz. Vùng ấy là chỗ bản trộn nghe đục.

10.3.2 Đo độ che lấp

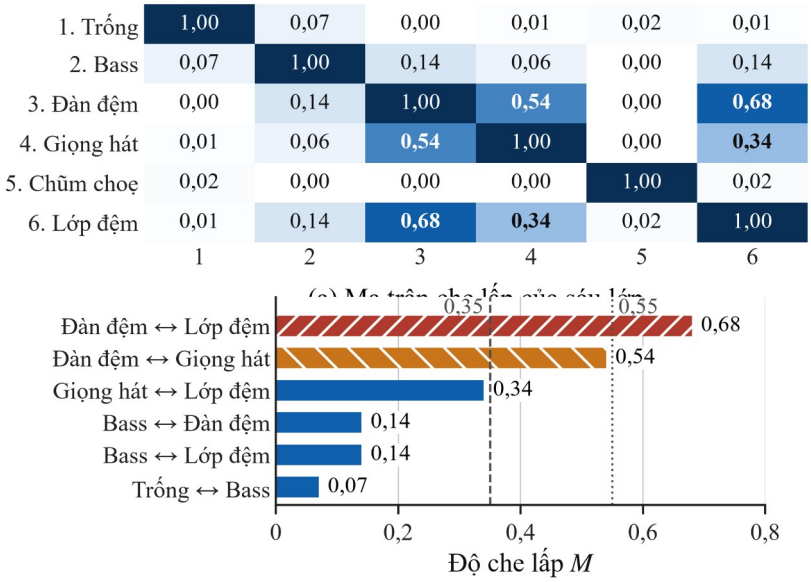
Nhìn hình thì thấy được, nhưng nhìn hình không xếp thứ tự được. Cần một con số. Cách đo dùng trong bài chia phổ thành hai mươi bốn dải theo thang loga, tính tỉ lệ năng lượng của từng lớp trong từng dải, rồi cộng phần chồng nhau.

$$M = \sum_{k=1}^{K} \min\left(p_k^A,\, p_k^B\right), \qquad p_k = \frac{E_k}{\sum_j E_j} \tag{10.4}$$

Trong đó p_k là tỉ lệ năng lượng của một lớp ở dải thứ k — tổng của chúng bằng 1 — còn M là độ che lấp, chạy từ 0 tới 1. Bằng 0 nghĩa là hai lớp không đụng nhau tí nào; bằng 1 nghĩa là chúng chiếm đúng cùng một vùng phổ.

Liệt kê 10.2 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())
```



(b) Sáu cặp xung đột nhiều nhất

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

Bảng 10.6 Ba cặp xung đột nhiều nhất trong sáu lớp của bài

Cặp lớp	Độ che lấp	Đọc ra sao
Đàn đệm ↔ Lớp đệm	0,68	Cặp bất lợi nhất. Hai lớp gần như trùng vùng phổ; đây là việc phải làm trước tiên
Đàn đệm ↔ Giọng hát	0,54	Đáng lo vì một trong hai là giọng hát — lớp không được phép bị che
Giọng hát ↔ Lớp đệm	0,34	Còn chấp nhận được; xử lý sau khi đã tách xong cặp trên

Một ngưỡng thực dụng

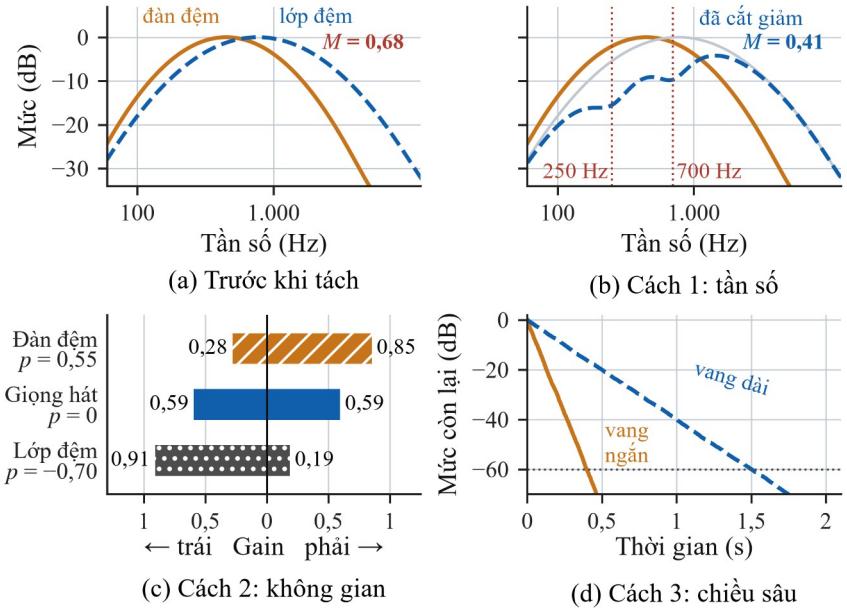
Dưới 0,35: hai lớp tự tách, không phải làm gì.

Từ 0,35 tới 0,55: nghe kỹ mới thấy; xử lý nếu một trong hai lớp là lớp quan trọng.

Trên 0,55: chắc chắn nghe thấy đục; phải tách.

Ngưỡng này rút ra từ cách đo ở trên với hai mươi bốn dải. Đổi số dải thì con số đổi theo, nên đừng mang ngưỡng này sang một cách đo khác.

10.3.3 Ba cách tách hai lớp ra khỏi nhau



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

Bảng 10.7 Ba cách tách và cái giá của từng cách

Cách	Công cụ	Được gì	Mất gì
1 — Tần số	Bộ EQ của mục 9.4	Hiệu quả nhất, đo được ngay bằng độ che lấp	Đổi màu âm của lớp bị cắt giảm
2 — Không gian	Pan của mục 10.4 và Haas của mục 10.5	Không đụng gì tới phổ	Chỉ dùng được cho lớp không bắt buộc nằm giữa; và có thể sai khi cộng về mono
3 — Chiều sâu	Bộ vang của mục 9.7	Tai tách được hai lớp dù phổ vẫn chồng	Thêm đuôi vang là thêm năng lượng, phải tính lại headroom

Ba cách này cộng dồn được chứ không loại trừ nhau. Trong hình, cách thứ nhất áp lên chính cặp bất lợi nhất: cắt giảm hai chỗ trên lớp đệm — 250 hertz hạ 9 decibel với Q bằng 0,9, và 700 hertz hạ 8 decibel với Q bằng 1,1 — đưa độ che lấp từ 0,68 xuống 0,41.

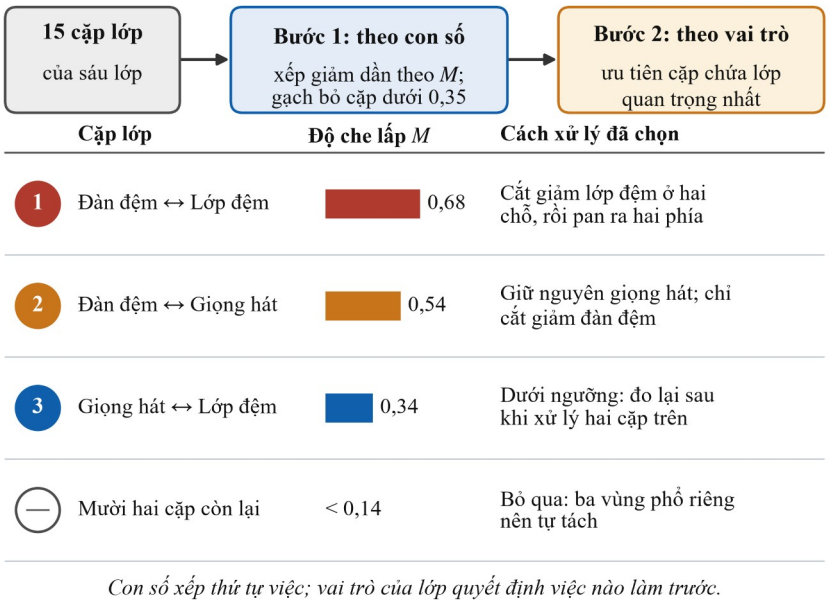
Vì sao cắt giảm lớp này thay vì nâng lớp kia

Hai cách nghe như tương đương: muốn giọng hát nổi lên thì hoặc nâng giọng hát, hoặc cắt giảm lớp đệm ở đúng vùng ấy.

Nhưng chúng không tương đương về headroom. Nâng thì cột đỉnh cao lên và có thể vượt trần; cắt giảm thì cột đỉnh hạ xuống.

Đây chính là lập luận "cắt an toàn hơn nâng" của mục 9.4.6, áp vào bài toán nhiều lớp.

10.3.4 Thứ tự ưu tiên khi có nhiều cặp phải xử lý



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

Một bản trộn sáu lớp có mười lăm cặp. Không xử lý hết được, và cũng không cần. Quy tắc xếp thứ tự gồm hai bước.

- **Bước một — theo con số.** Xếp mười lăm cặp theo độ che lấp giảm dần, gạch bỏ mọi cặp dưới 0,35.
- **Bước hai — theo vai trò.** Trong số còn lại, ưu tiên cặp nào có chứa lớp quan trọng nhất của bản nhạc — thường là giọng hát hoặc nhạc cụ chính.

Một lỗi hay gặp ở bước hai

Người mới thường xử lý cặp có con số cao nhất trước, bất kể cặp ấy gồm những lớp nào. Nhưng hai lớp nền che nhau 0,68 hại ít hơn một lớp nền che giọng hát 0,54, vì tai người nghe nhạc bám vào giọng hát.

Con số xếp thứ tự việc; vai trò của lớp quyết định việc nào làm trước.

Áp hai bước ấy vào sáu lớp của bài thì ra danh sách việc dưới đây. Chú ý cột cuối: cách xử lý không giống nhau cho mọi cặp, vì vai trò của các lớp trong cặp khác nhau.

Bảng 10.8 Danh sách việc rút ra từ hai bước xếp thứ tự

Thứ tự	Cặp lớp	Che lấp	Cách xử lý đã chọn và lý do
1	Đàn đệm ↔ Lớp đệm	0,68	Cả hai đều là lớp nền nên đều cắt giảm được. Cắt giảm lớp đệm ở hai chỗ, rồi pan hai lớp ra hai phía — dùng cả cách tần số lẫn cách không gian
2	Đàn đệm ↔ Giọng hát	0,54	Giọng hát *không* được đụng tới. Chỉ cắt giảm đàn đệm, và chỉ ở đúng vùng giọng hát trội
3	Giọng hát ↔ Lớp đệm	0,34	Dưới ngưỡng, và sau khi xử lý hai cặp trên thì con số này tự giảm theo. Đo lại rồi mới quyết định có làm gì không
—	Mười hai cặp còn lại	dưới 0,14	Bỏ qua. Trống, bass và chũm chọe nằm ở ba vùng phổ riêng nên chúng tự tách

10.4 Panning — đặt lớp vào chỗ giữa hai loa

10.4.1 Pan là gì, và nó khác gì với chỉnh mức

Pan là chia một lớp mono thành hai kênh với hai gain khác nhau. Gọi p là vị trí pan, chạy từ −1 (trái hết) tới +1 (phải hết), thì luật pan là quy tắc sinh ra cặp gain (g_L, g_R) từ p.

Câu hỏi thật sự không phải "chia thế nào" mà là "chia sao cho khi kéo p từ trái sang phải thì nguồn không thay đổi độ to". Câu hỏi ấy không có lời giải đúng, và biết nó không có lời giải đúng mới là điều đáng học ở mục này.

10.4.2 Luật thứ nhất — giữ biên độ

$$g_L = 1 - q,\ g_R = q;\qquad q = \frac{p+1}{2},\ p \in [-1,\ 1] \tag{10.5}$$

Luật đơn giản nhất: hai gain cộng lại luôn bằng 1. Ở chính giữa, mỗi kênh nhận 0,5.

Ưu điểm của luật này lộ ra khi cộng về mono: hai kênh cộng thẳng lại cho đúng tín hiệu gốc, không hơn không kém, ở mọi vị trí pan. Nhược điểm lộ ra khi nghe qua hai loa: ở chính giữa, tổng *công suất* chỉ còn một nửa, tức là hụt 3,01 decibel. Kéo một nhạc cụ vào giữa là nó tụt xuống.

10.4.3 Luật thứ hai — giữ công suất

$$g_L = \cos\frac{\pi q}{2},\quad g_R = \sin\frac{\pi q}{2};\qquad g_L^2 + g_R^2 = 1 \tag{10.6}$$

Luật này dùng sin và cosin nên tổng bình phương hai gain luôn bằng 1. Nghe qua hai loa thì độ to phẳng tuyệt đối ở mọi vị trí pan — đúng cái mà luật trước không làm được.

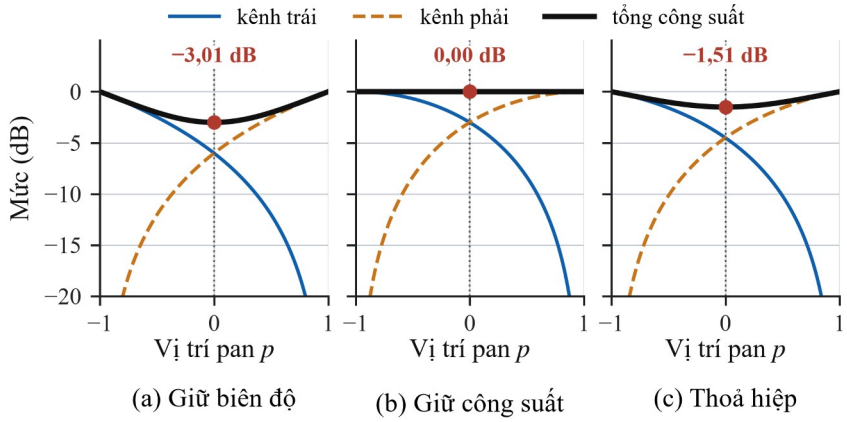
Nhưng khi cộng về mono thì sai theo chiều ngược lại: ở chính giữa, mỗi kênh nhận 0,707, cộng thẳng lại được 1,414 — vống lên đúng 3,01 decibel. Một bản trộn dùng luật này, khi phát trên loa điện thoại, sẽ có mọi thứ ở giữa to hơn 3 decibel so với ý định.

10.4.4 Vì sao không luật nào đúng cả hai đầu

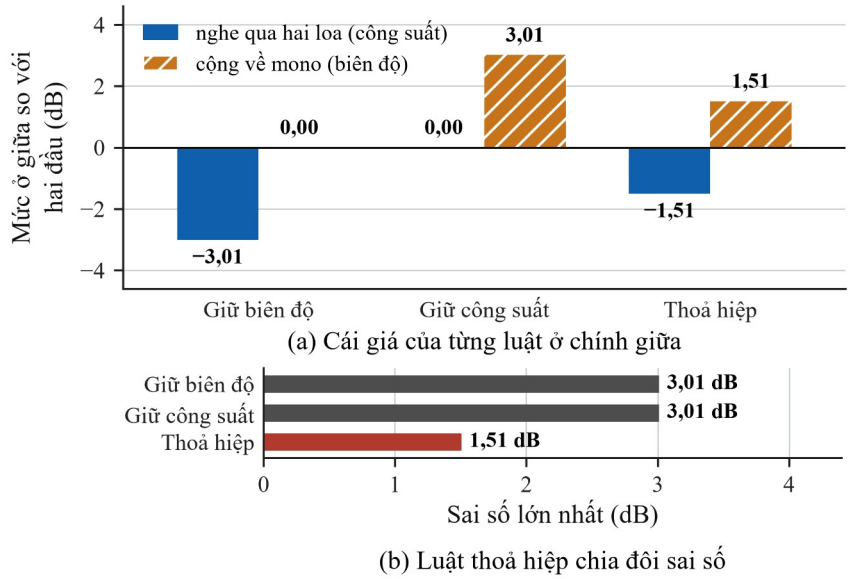
$$L_{\text{cs}} = 10\log_{10}\left(g_L^2 + g_R^2\right);\qquad L_{\text{mono}} = 20\log_{10}\left(g_L + g_R\right) \tag{10.7}$$

Hai dòng của công thức (10.7) là hai phép đo khác nhau trên cùng một cặp gain. Dòng trên là mức nghe được qua hai loa: hai kênh đi tới tai theo hai đường khác nhau nên *công suất* cộng lại. Dòng dưới là mức sau khi cộng về mono: hai kênh cộng thẳng trong mạch nên *biên độ* cộng lại.

Không có cặp gain nào giữ được cả hai dòng phẳng cùng lúc. Đó là một sự thật toán học, không phải hạn chế của thiết bị.



Hình 10.11 Ba luật pan nhìn theo đường gain



(a) Cái giá của từng luật ở chính giữa

(b) Luật thoả hiệp chia đôi sai số

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

Bảng 10.9 Ba luật pan và cái giá của mỗi luật

Luật	Hụt ở giữa khi nghe loa	Vống khi cộng mono	Dùng khi
Giữ biên độ	−3,01 dB	0,00 dB	Bản trộn chắc chắn sẽ bị nghe ở chế độ mono — truyền thanh, thông báo công cộng
Giữ công suất	0,00 dB	+3,01 dB	Bản trộn chỉ nghe qua hai loa hoặc tai nghe — trò chơi, phim
Thoả hiệp	−1,51 dB	+1,51 dB	Mặc định của gần như mọi bàn trộn: chia đôi sai số cho cả hai đầu

Liệt kê 10.3 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
```

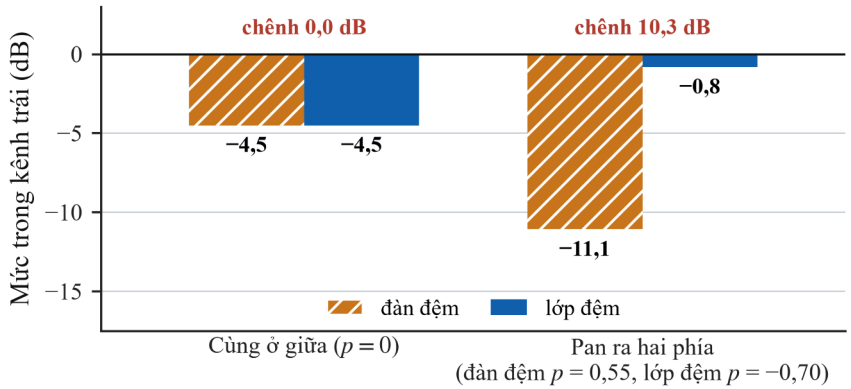
Luật thoả hiệp được dựng ra sao

Nó là trung bình nhân của hai luật kia: lấy căn bậc hai của tích hai cặp gain.

Vì trung bình nhân trong miền hệ số tương ứng với trung bình cộng trong miền decibel, sai số của nó chia đôi đều cho cả hai đầu — đúng 1,51 decibel mỗi bên, bằng nửa của 3,01.

Đây là ví dụ tốt về một quyết định kỹ thuật không có đáp án đúng: lời giải hay nhất là lời giải chia đều thiệt hại và *nói rõ đang chia đều* thiệt hại.

10.4.5 Pan là cách tách lớp không đụng tới phổ



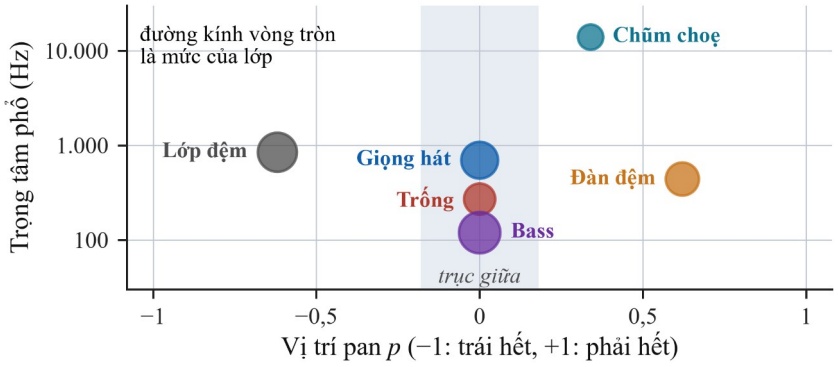
(a) Chênh mức trong một kênh, tai dùng để tách hai lớp

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

Quay lại cặp bất lợi nhất của mục 10.3 — đàn đệm và lớp đệm. Khi cả hai cùng nằm giữa thì trong mỗi kênh chúng ngang mức nhau, nên tai không có gì để bám vào. Pan chúng ra hai phía thì trong kênh trái, một lớp to hơn lớp kia hơn mười decibel, và tai dùng đúng chênh lệch ấy để tách hai lớp.

Điều đáng chú ý là phép này *không* đụng gì tới phổ: độ che lấp đo trên tín hiệu mono vẫn nguyên 0,68. Nó tách theo một chiều khác hẳn.

10.4.6 Bản đồ đặt chỗ cho cả bản trộn



(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

Hình trên gói toàn bộ quyết định đặt chỗ của một bản trộn vào một chỗ. Trục ngang là vị trí pan, trục dọc là trọng tâm phổ của từng lớp, đường kính vòng tròn là mức của lớp.

Bảng 10.10 Bốn quy ước đặt chỗ và lý do của từng quy ước

Quy ước	Lý do
Trống, bass và giọng hát để ở trục giữa	Tai phân biệt hướng ở chính diện nhạy gấp mười lần ở hai bên — xem mục 10.6.4; và vùng trầm đẩy ra một bên sẽ làm lệch headroom giữa hai kênh
Cặp xung đột nhất tách sang hai phía	Đây là cách tách thứ hai của mục 10.3.3, dùng thêm chứ không thay cho cách tần số
Không để quá hai lớp ở đúng chính giữa	Càng nhiều lớp ở giữa thì hai kênh càng giống nhau, và độ rộng stereo càng về 0
Lớp cao tần có thể pan mạnh tay hơn	Ở dải cao, tai định vị bằng chênh mức nên pan có hiệu lực rõ; thêm nữa mất một phần cao tần khi cộng mono ít hại hơn mất trầm

Bảng 10.11 Bốn con số đo trên bản trộn trước và sau khi đặt chỗ

Con số	Trước	Sau	Đọc ra sao
Độ rộng stereo	0,000	0,124	Từ mono thật thành có bề rộng. Khoảng 0,10 tới 0,60 là vùng lành mạnh
Tách cặp bất lợi nhất trong một kênh	0,0 dB	11,4 dB	Đây là thứ tai dùng để phân biệt hai lớp
Gợn phổ khi cộng về mono	0,00 dB	0,32 dB	Vẫn dưới nửa decibel — bản trộn an toàn khi về mono
Đỉnh của bản trộn	−8,47 dBFS	−8,03 dBFS	Gần như không đổi, nên headroom đã chừa ở mục 10.2 vẫn đủ

Một điều bất ngờ ở con số cuối

Nhiều người nghĩ pan các lớp ra hai bên thì đỉnh phải *thấp* đi, vì mỗi kênh nhận ít lớp hơn.

Đo thì thấy ngược lại một chút: đỉnh nhích lên 0,44 decibel. Lý do là luật thoả hiệp cho gain gần 0,94 ở một kênh khi pan hết cỡ, trong khi ở giữa mỗi kênh chỉ nhận 0,595.

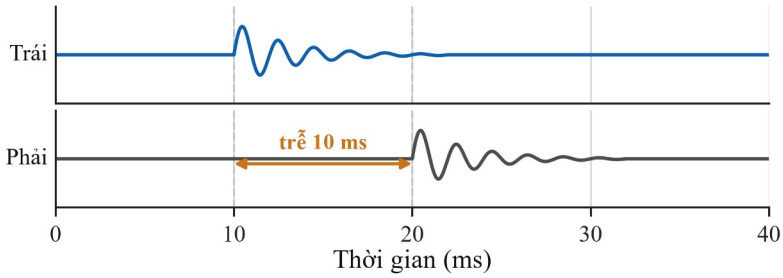
Chênh lệch nhỏ nên không đáng lo, nhưng nó nhắc một quy tắc: đo lại đỉnh *sau* khi đặt chỗ, đừng cho rằng pan luôn làm bản trộn nhỏ đi.

10.5 Hiệu ứng Haas — làm rộng bằng thời gian

10.5.1 Một điều đáng ngạc nhiên về tai người

Lấy một lớp mono, cho vào cả hai kênh với *cùng* một mức, rồi trễ một kênh đi mười mili giây. Về mặt tín hiệu, hai kênh chỉ khác nhau ở thời điểm.

Tai không nghe ra hai tiếng. Nó nghe ra *một* tiếng, đến từ phía kênh sớm hơn, và tiếng ấy rộng hơn hẳn so với khi hai kênh hoàn toàn giống nhau. Đó là hiệu ứng Haas, còn gọi là hiệu ứng đi trước.



(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

Cơ chế là một cơ chế sinh tồn: trong một căn phòng, âm trực tiếp tới trước, phản xạ tới sau. Nếu tai coi mỗi phản xạ là một nguồn riêng thì mọi phòng đều nghe thành hỗn loạn. Bộ não vì thế gộp những thứ tới trong khoảng vài chục mili giây đầu vào cùng một sự kiện, và lấy *hướng* của cái tới trước làm hướng của cả cụm.

10.5.2 Bảng năm mốc — cùng một kỹ thuật, năm kết quả

Bảng 10.12 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. Không còn định vị theo trễ mà theo màu âm — đây là lỗi
1 – 5 ms	Nguồn dịch hẳn về phía kênh sớm	Vùng làm rộng mạnh tay; tương thích mono kém nhất
5 – 35 ms	Vẫn *một* tiếng, đến từ phía kênh sớm	Vùng Haas đúng nghĩa — làm rộng mà vẫn tự nhiên
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 nữa — đó là slapback của mục 8.7

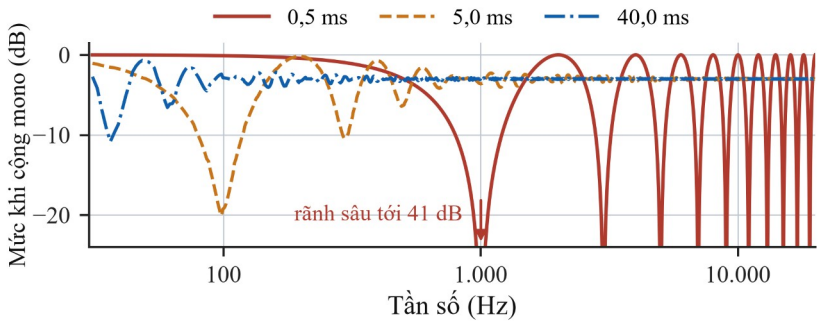
Bảng này là thứ quyết định: cùng một dòng mã, chỉ đổi con số mili giây, mà ra năm kết quả khác hẳn nhau. Đó là lý do hiệu ứng này hay bị dùng sai — người dùng chỉnh núm cho tới khi "nghe rộng" mà không biết đã đi vào vùng nào.

10.5.3 Cái giá của Haas — và nó chỉ hiện ra ở chế độ mono

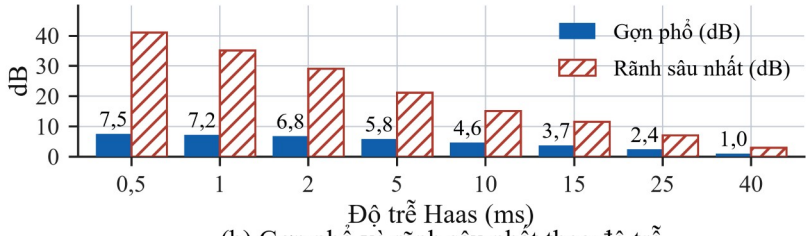
Chừng nào còn nghe hai kênh riêng thì Haas không mất gì. Nhưng khi ai đó cộng hai kênh về mono — điện thoại, loa Bluetooth đơn, hệ truyền thanh — thì phép cộng ấy chính là bộ lọc lược của mục 8.5.

$$m[n] = \frac{x[n] + x[n-D]}{2}; \qquad f_{\text{rãnh}} = \frac{(2k+1)f_s}{2D} \tag{10.8}$$

Dòng thứ nhất là phép cộng về mono; dòng thứ hai là vị trí các rãnh của bộ lọc lược sinh ra, với D là số mẫu trễ và k chạy từ 0 trở lên. Rãnh đầu tiên nằm ở tần số bằng nửa nghịch đảo của thời gian 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

Bảng 10.13 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 nằm ngay giữa vùng nghe, và sâu tới 40 decibel. Đây là chỗ nguy nhất
5 ms	100 Hz	vài decibel	Rãnh dày hơn nên tai coi chúng như một sự đổi màu nhẹ
15 ms	33,3 Hz	nhỏ	Rãnh dày tới mức nằm gọn trong một dải phân giải của tai
40 ms	12,5 Hz	1,00 dB	Gần như không đo được. Trễ *dài* lại an toàn hơn trễ ngắn

Chiều của quan hệ này ngược với trực giác

Trực giác nói: trễ càng dài thì càng "sai lệch" nhiều, nên càng nguy.

Số đo nói ngược lại. Trễ *ngắn* mới nguy, vì rãnh của nó nằm thưa và rơi đúng vào giữa vùng nghe. Ở 0,5 mili giây, gợn tới 7,5 decibel với rãnh sâu hơn 40 decibel; ở 40 mili giây, gợn chỉ còn 1,0 decibel.

Vì sao? Vì rãnh cách nhau đúng nghịch đảo của thời gian trễ. Trễ dài cho rãnh dày, và rãnh dày thì tai gộp lại thành một mức trung bình phẳng.

10.5.4 Bốn cách làm rộng và cách chọn giữa chúng

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

Haas chỉ là một trong bốn cách. Bảng trong hình xếp cả bốn cạnh nhau để thấy rõ chúng không thay thế nhau: mỗi cách trả một cái giá khác.

- **Pan — đổi mức.** Cộng mono không sao, vị trí rõ ràng; nhưng nguồn dịch hẳn sang một bên và mất chỗ ở giữa.
- **Haas — đổi thời điểm.** Rất rộng mà vẫn nghe gần giữa; nhưng cộng mono thành lọc lược.
- **Hai bản thu.** Rộng thật và tương thích mono tốt nhất, vì hai bản thu độc lập nên hệ số tương quan gần 0; nhưng phải thu lại, không dùng được cho tư liệu có sẵn.
- **Vang ngắn khác nhau hai bên.** Rộng và tự nhiên, không đụng tới vị trí; nhưng đắt và dễ làm bản trộn nghe xa xăm.

Một nút bấm bắt được cả bốn lỗi

Trước khi giao bất kỳ bản trộn nào, bấm nút mono và nghe lại.

Nếu một lớp mỏng đi hoặc biến mất thì có Haas trễ ngắn, hoặc có lớp bị đảo cực như ở mục 10.2.2.

Nếu mọi thứ ở giữa vống lên thì luật pan đang dùng là luật giữ công suất.

Cách đo thay cho tai: hàm gon_pho_khi_ve_mono_db trả về độ lệch chuẩn của đáp ứng và độ sâu của rãnh sâu nhất. Dưới 3 decibel là đạt.

10.6 ITD và ILD — tai người định vị bằng gì

10.6.1 Hai manh mối, và một câu hỏi cũ hơn trăm năm

Pan và Haas là hai kỹ thuật của người trộn nhạc. Chúng hiệu quả vì chúng mô phỏng hai thứ mà tai người thật sự dùng để định vị: chênh lệch thời gian giữa hai tai, và chênh lệch mức giữa hai tai. Mục này đo cả hai.

Bảng 10.14 Ba manh mối định vị và vùng tần số của từng manh mối

Manh mối	Làm việc ở đâu	Cơ chế
ITD — chênh thời gian	Dưới ~1,5 kHz	Tai so *pha* của hai bên; lớn nhất chừng 660 micro giây
ILD — chênh mức	Trên ~1,5 kHz	Đầu chắn sóng ngắn; có thể lên tới 15 – 20 decibel
Phổ vành tai	Trên ~5 kHz	Vành tai tạo rãnh phổ khác nhau theo độ *cao* — đây là thứ giải quyết được nón lẫn lộn, xem mục 10.7

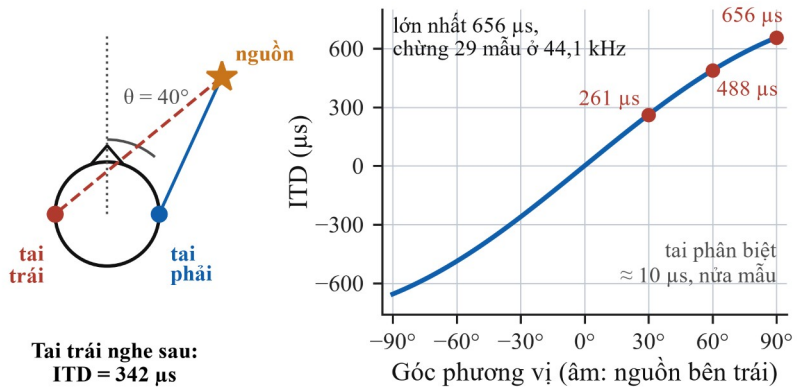
10.6.2 ITD — chênh lệch thời gian giữa hai tai

Sóng tới từ bên phải chạm tai phải trước, rồi phải đi vòng quanh đầu mới tới tai trái. Woodworth đưa ra công thức xấp xỉ cho quãng đường vòng ấy, với a là bán kính đầu và c là tốc độ âm.

$$\mathrm{ITD}(\theta) = \frac{a}{c}(\theta + \sin\theta) \tag{10.9}$$

Với bán kính đầu người trưởng thành lấy tròn 8,75 centimet và tốc độ âm 343 mét trên giây, giá trị lớn nhất — khi nguồn ở đúng bên hông — bằng:

$$\mathrm{ITD}_{\max} = \frac{a}{c}\left(\frac{\pi}{2} + 1\right) \approx 656\,\mu\mathrm{s} \tag{10.10}$$



(a) Hình học (b) ITD theo góc (Woodworth)

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

Con số 656 micro giây đáng dừng lại một chút

Toàn bộ khả năng định vị trái phải của con người nằm gọn trong một khoảng 656 micro giây — chưa tới ba mươi mẫu ở 44,1 kilohertz.

Trong khoảng bé tí ấy, tai phân biệt được chênh lệch chừng mười micro giây, tức nửa mẫu.

Đó là lý do phép trễ phân số của mục 9.2 không phải chuyện lý thuyết: muốn đặt nguồn chính xác thì phải trễ được những lượng nhỏ hơn một mẫu. Trễ theo mẫu nguyên cho độ phân giải thô gấp đôi độ phân giải của tai.

Liệt kê 10.4 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)))

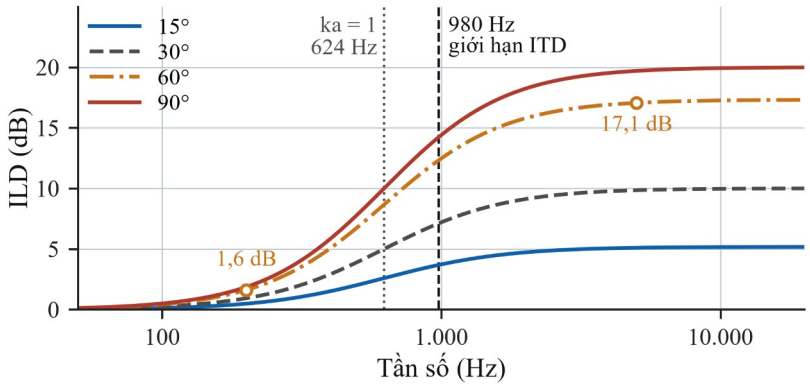
def goc_tu_itd(itd_s, ban_kinh=BAN_KINH_DAU, c=TOC_DO_AM):
    """Phương trình Woodworth không có nghiệm kín — dò nhị phân."""
    dau = 1.0 if itd_s >= 0 else -1.0
    lo, hi = 0.0, 90.0
    for _ in range(60):
        gi = (lo + hi) / 2
        if itd_giay(gi, ban_kinh, c) < abs(itd_s):
          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

10.6.3 ILD — chênh lệch mức, và vì sao nó chỉ có ở dải cao

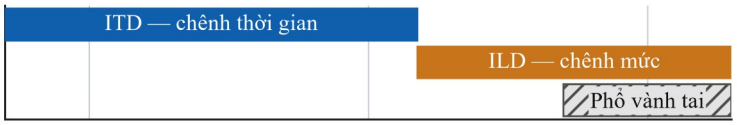
Đầu người chắn sóng, nhưng chỉ chắn được sóng có bước sóng đủ ngắn. Đại lượng quyết định là tích của số sóng k với bán kính đầu a.

$$ka = \frac{2\pi f a}{c}\,; \qquad \mathrm{ILD} \approx 20\,\frac{(ka)^2}{1+(ka)^2}\sin\theta \ \text{[dB]} \tag{10.11}$$

Khi ka nhỏ hơn 1, bước sóng dài hơn đầu nên sóng vòng qua được và chênh mức gần bằng không. Khi ka lớn hơn 1, đầu chắn hẳn và số hạng bóng đầu bão hoà về 1. Chỗ chuyển tiếp — ka bằng đúng 1 — nằm ở 624 hertz.



(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

Ở góc 60 độ, chênh mức hai tai chỉ 1,6 decibel tại 200 hertz nhưng lên tới 17,1 decibel tại 5 kilohertz. Cùng một nguồn, cùng một hướng, mà hai dải tần cho hai con số cách nhau hơn mười lần.

10.6.4 Thuyết song công và ranh giới của nó

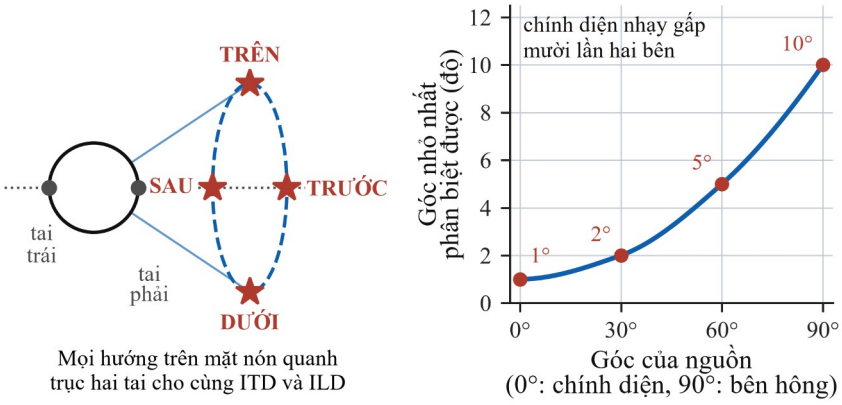
Rayleigh gộp hai quan sát trên thành thuyết song công: dải trầm định vị bằng ITD, dải cao định vị bằng ILD. Ranh giới không phải do sở thích mà do một giới hạn vật lý.

$$f_{\text{gh}} = \frac{c}{4a} \approx 980 \text{ Hz} \tag{10.12}$$

Trên tần số này, nửa bước sóng ngắn hơn khoảng cách hai tai, nên chênh pha vượt 180 độ và tai không còn biết tai nào sớm hơn — pha bị lẫn vòng. Đúng lúc ITD hết dùng được thì ILD bắt đầu đủ lớn, nên hai manh mối bù cho nhau.

> **Hai con số gần nhau mà khác nghĩa**
>
> 624 hertz là chỗ ka bằng 1 — nơi bóng đầu *bắt đầu* có hiệu lực, tức là nơi ILD bắt đầu lớn dần.
>
> 980 hertz là chỗ nửa bước sóng bằng đường kính đầu — nơi ITD *hết* dùng được vì pha lẫn vòng.
>
> Cả hai đều nằm quanh một kilohertz, và đó không phải trùng hợp: cả hai đều tính từ cùng một bán kính đầu. Chính vì chúng gần nhau mà hai manh mối nối được vào nhau không để lại khoảng trống.

10.6.5 Chỗ mà hai manh mối cùng bất lực



(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

Hai manh mối trên đủ để biết nguồn ở bên trái hay bên phải, nhưng không đủ để biết nó ở trên hay ở dưới, ở trước hay ở sau. Lý do hình học: mọi hướng nằm trên một mặt nón quanh trục hai tai đều cho cùng một ITD và cùng một ILD. Đó là nón lẫn lộn.

Hình (b) của cùng hình ấy giải thích một quy ước trộn nhạc đã dùng ở mục 10.4.6: góc nhỏ nhất mà tai còn phân biệt được là 1 độ ở chính diện nhưng tới 10 độ ở bên hông. Tai nhạy hướng ở chính diện gấp mười lần ở hai bên — đó là lý do những lớp quan trọng nhất được để ở giữa.

Thực tế người ta ít khi lẫn trước với sau, vì bộ não còn ba đường thoát nữa. Bảng dưới đây xếp cả bốn, và nó cũng là bảng nói rõ đường nào lập trình được, đường nào không.

Bảng 10.15 Bốn cách giải quyết nón lẫn lộn

Cách	Cơ chế	Lập trình được không
Dấu vết phổ của vành tai	Vành tai tạo rãnh phổ ở vị trí đổi theo độ cao — mục 10.7.3	Được, nhưng phải có HRTF đo thật, và HRTF của mỗi người một khác
Chuyển động đầu	Quay đầu vài độ thì ITD và ILD đổi theo, và chiều đổi khác nhau giữa nguồn trước với nguồn sau	Được, nếu thiết bị có cảm biến hướng. Đây là cách hiệu quả nhất trong thực tế ảo
Thông tin thị giác	Mắt thấy nguồn ở đâu thì tai tin theo — hiệu ứng người bụng nói	Được, và gần như miễn phí trong trò chơi: chỉ cần vẽ nguồn ở đúng chỗ
Hiểu biết về bối cảnh	Tiếng bước chân thường ở dưới, tiếng chim thường ở trên	Không lập trình được, nhưng nên biết là nó có — nó che bớt sai số của ba cách trên

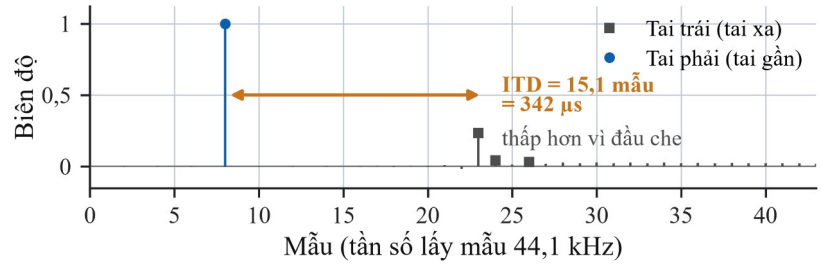
Cách thứ nhất là nội dung của mục tiếp theo. Ba cách còn lại giải thích vì sao một hệ binaural có thêm cảm biến hướng đầu lại nghe thuyết phục hơn hẳn một hệ chỉ có HRTF tốt.

10.7 Binaural và HRTF

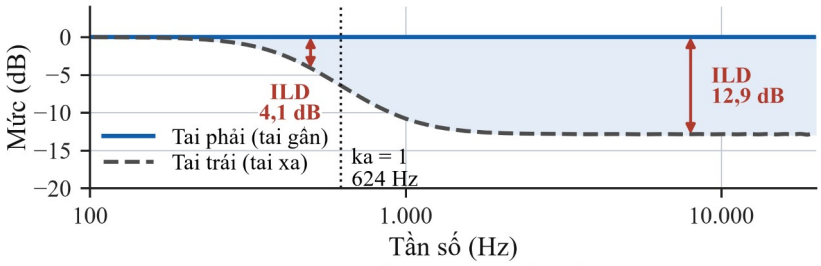
10.7.1 Từ hai manh mối tới một bộ lọc

Mục trước đã có đủ vật liệu để đặt một nguồn vào một hướng: trễ tai xa đúng ITD của công thức (10.9), rồi hạ mức tai xa theo công thức (10.11). Hai dòng, và nguồn nghe ra ở đúng bên.

Nhưng tai người không chỉ nhận hai con số. Nó nhận một đáp ứng xung đầy đủ cho mỗi tai — đường đi từ nguồn tới màng nhĩ, đã bị đầu, vai và vành tai làm biến dạng. Cặp đáp ứng xung ấy gọi là HRIR; nhìn theo tần số thì gọi là 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

$$y_L[n] = \left(h_L^{\theta,\varphi} * x\right)[n]\,; \qquad y_R[n] = \left(h_R^{\theta,\varphi} * x\right)[n] \tag{10.13}$$

Công thức (10.13) là toàn bộ cách dùng: tích chập tín hiệu mono với đáp ứng xung của tai trái để ra kênh trái, và với đáp ứng xung của tai phải để ra kênh phải. Chỉ số theta và phi là phương vị và độ cao — đổi hướng nghĩa là đổi cặp đáp ứng xung.

HRIR và HRTF là một thứ

HRIR là cặp đáp ứng xung, sống trong miền thời gian.

HRTF là cùng cặp ấy sau phép biến đổi Fourier, sống trong miền tần số.

Dùng tên nào là tuỳ chỗ: nói về cài đặt thì gọi HRIR vì chúng ta tích chập với nó; nói về đặc tính nghe được thì gọi HRTF vì chúng ta đọc rãnh và đỉnh nhô trên đó.

Đây đúng là quan hệ giữa đáp ứng xung và đáp ứng tần số của mục 9.8, chỉ khác là bây giờ có hai cái, một cho mỗi tai.

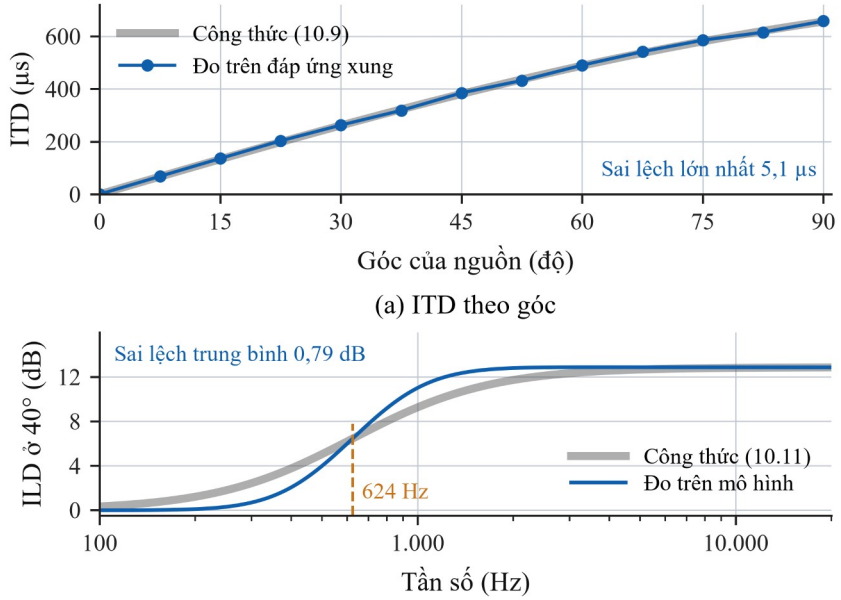
10.7.2 Một mô hình binaural đơn giản, và nó đúng tới đâu

Liệt kê 10.5 Đặt một 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])
```

Hàm này dùng lại nguyên vẹn hai khối của chương 9: trễ phân số của mục 9.2 để đặt được ITD nhỏ hơn một mẫu, và bộ lọc kệ của mục 9.3 để làm bóng đầu. Tần số cắt của bộ lọc kệ không chọn tuỳ ý — nó đặt đúng ở chỗ ka bằng 1, tức điểm uốn của công thức (10.11).



(a) ITD theo góc

(b) ILD theo tần số

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

Bảng 10.16 Kết quả phép kiểm đặt rồi đo

Đại lượng	Cách kiểm	Sai lệch	Kết luận
ITD	Đo khoảng cách hai đỉnh trên đáp ứng xung, nội suy parabol để có phần lẻ của mẫu	Lớn nhất 4,8 µs trên cả dải góc	Cài đặt khớp công thức (10.9) — dưới một phần trăm mẫu
ILD	Lấy hiệu hai đường |H(f)| của cặp đáp ứng xung, so với công thức	Trung bình 0,79 dB	Chỉ khớp *đại khái*: một bộ lọc kệ bậc hai không phải đường cong bóng đầu

Đây là phép kiểm cài đặt, không phải kiểm mô hình

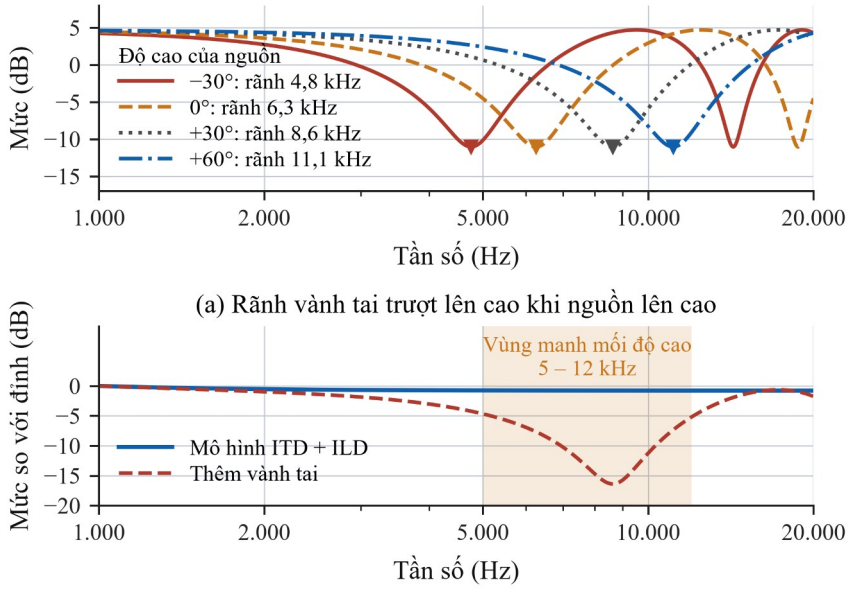
Phép kiểm này chỉ nói rằng đoạn mã làm đúng điều công thức yêu cầu. Nó *không* nói rằng công thức mô tả đúng tai người.

Phân biệt hai chuyện ấy là một thói quen nghề nghiệp: một mô hình sai vẫn có thể được cài đặt hoàn hảo.

Sai số 0,79 decibel của ILD nghe được: nguồn hơi lệch chỗ ở dải 500 hertz tới 2 kilohertz. Muốn khớp hơn thì phải dùng HRTF đo thật, không dùng mô hình nữa.

10.7.3 Chỗ trống lớn nhất của mô hình đơn giản

Mô hình trên có ITD và có ILD, tức là có hai trong ba manh mối của bảng ở mục 10.6.1. Manh mối thứ ba — dấu vết phổ của vành tai — hoàn toàn không có, và đó là lý do nguồn dựng bằng mô hình ấy nghe được trái phải nhưng không biết trên dưới.



(a) Rãnh vành tai trượt lên cao khi nguồn lên cao

(b) Chỗ trống của mô hình đơn giản

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

Cơ chế của manh mối thứ ba đơn giản một cách đáng ngạc nhiên. Âm tới thẳng vào ống tai, đồng thời một phần dội lại từ gờ vành tai rồi mới vào. Hai đường

ấy cộng nhau chính là bộ lọc lược của mục 8.5, với rãnh đầu tiên ở nửa nghịch đảo của thời gian trễ.

Bảng 10.17 Rãnh vành tai trượt theo độ cao — mô hình phản xạ đơn giản hoá

Độ cao của nguồn	Trễ phản xạ	Rãnh đầu tiên
−30° (dưới tầm tai)	105 µs	4,8 kHz
0° (ngang tầm tai)	80 µs	6,3 kHz
+30°	58 µs	8,6 kHz
+60° (gần đỉnh đầu)	45 µs	11,1 kHz

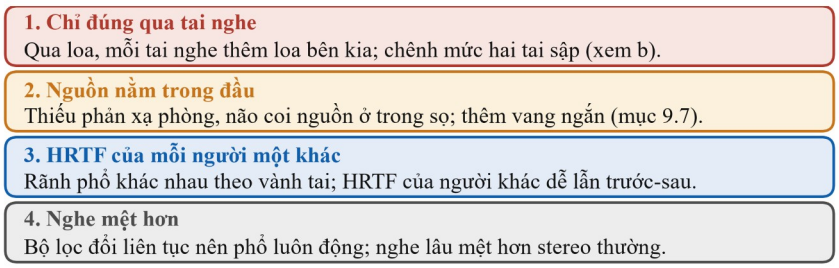
Một lưu ý về nguồn gốc các con số trên

Bảng này dựng từ mô hình phản xạ vành tai *đơn giản hoá*, không phải số đo trên người thật.

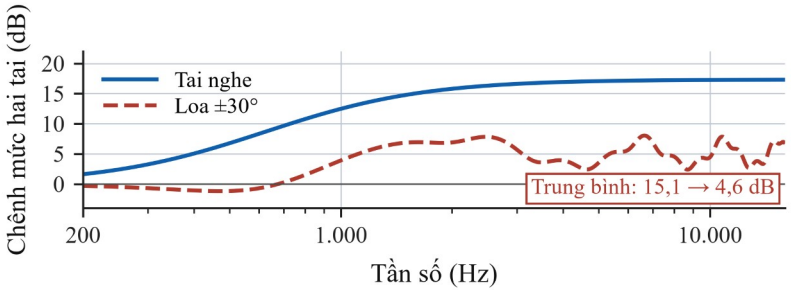
Nó chỉ đủ để giải thích vì sao rãnh trượt lên cao khi nguồn lên cao: nguồn càng cao thì đường phản xạ qua gờ vành tai càng ngắn.

HRTF đo thật phức tạp hơn nhiều — có nhiều rãnh, và vị trí của chúng khác nhau giữa từng người. Đó chính là giới hạn thứ ba ở mục sau.

10.7.4 Bốn giới hạn phải biết trước khi dùng binaural



(a) Bốn giới hạn phải biết trước khi dùng



(b) Nguồn đặt ở 60°, phát qua hai đường

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

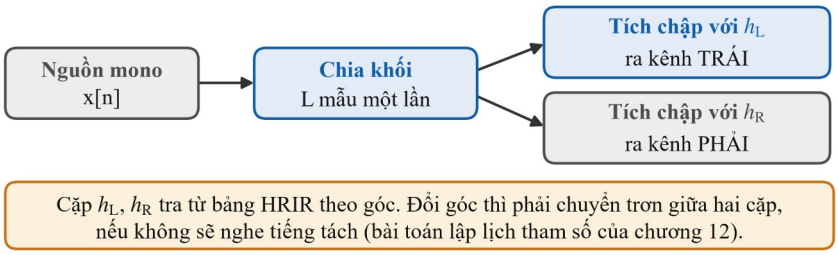
Bảng 10.18 Bốn giới hạn của âm thanh binaural

Giới hạn	Nội dung	Làm gì được
1 — Chỉ đúng qua tai nghe	Qua loa, mỗi tai nghe thêm loa bên kia. Chênh mức hai tai của một nguồn ở 60 độ sập từ 15,1 xuống 4,6 decibel	Ghi rõ "nghe bằng tai nghe"; hoặc dùng bộ khử xuyên âm, nhưng nó chỉ đúng ở một chỗ ngồi duy nhất
2 — Nguồn nằm trong đầu	Không có phản xạ phòng thì não coi nguồn là ở trong sọ chứ không ở ngoài	Thêm một bộ vang ngắn của mục 9.7 — chỉ vài phản xạ sớm đã đủ đẩy nguồn ra ngoài
3 — HRTF của mỗi người một khác	Vành tai khác nhau nên rãnh phổ nằm ở tần số khác nhau. Dùng HRTF của người khác thì hay lẫn trước với sau	Cho người dùng chọn trong vài bộ HRTF; hoặc thêm chuyển động đầu, vì đầu quay một chút là giải quyết được nón lẫn lộn
4 — Nghe mệt hơn	Bộ lọc thay đổi liên tục nên phổ luôn động	Đừng dùng binaural cho mọi lớp; giữ nó cho những nguồn cần định vị thật sự

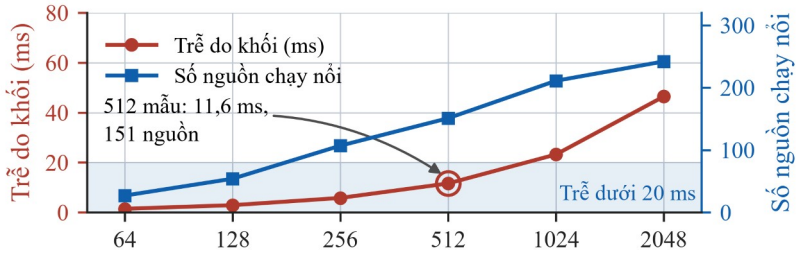
Giới hạn thứ nhất là giới hạn duy nhất trong bốn giới hạn trên đo được ngay bằng vài dòng mã. Mô phỏng cặp loa đặt ±30 độ: mỗi tai nghe thêm loa bên kia, trễ chừng 0,26 mili giây và nhỏ hơn 3 decibel. Chênh mức hai tai — thứ mang toàn bộ thông tin hướng ở dải cao — tụt từ 15,1 xuống 4,6 decibel, và ở dải trên 4 kilohertz thì gợn lên xuống chứ không còn là một con số.

> **Ba trong bốn giới hạn là lý do binaural ít dùng cho nhạc**
>
> Một bản nhạc phát hành phải nghe được trên mọi thứ: tai nghe, loa, điện thoại, xe hơi. Giới hạn thứ nhất loại ngay ba trong bốn thứ ấy.
>
> Trò chơi và thực tế ảo thì khác: người chơi gần như luôn đeo tai nghe, và có thêm chuyển động đầu để bù giới hạn thứ ba.
>
> Đó là lý do binaural là chuyện thường ngày trong lập trình trò chơi mà gần như vắng mặt trong sản xuất nhạc.

10.7.5 Giá phải trả để chạy thời gian thực



(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

Đặt một nguồn vào một hướng là hai phép tích chập, mỗi phép với một đáp ứng xung 256 mẫu. Cài đặt dùng đúng hàm tích chập theo khối của mục 9.8, và kích thước khối là tham số duy nhất phải chọn.

Bảng 10.19 Chọn kích thước khối — đánh đổi giữa hai thứ

Khối	Trễ do khối	Số nguồn chạy nổi	Nhận xét
64 mẫu	1,5 ms	ít nhất	Trễ rất thấp nhưng chi phí mỗi mẫu cao — phù hợp nhạc cụ điện tử chơi trực tiếp
256 mẫu	5,8 ms	khá	Chỗ nhiều bộ máy trò chơi dừng lại
512 mẫu	11,6 ms	nhiều	Chỗ thường chọn: trễ vẫn dưới ngưỡng nhận biết, số nguồn đã dư cho một cảnh thường
1024 mẫu trở lên	23 ms trở lên	gần như không tăng thêm	Chỗ gãy: từ đây trễ tăng gấp đôi mỗi bước mà số nguồn hầu như đứng yên

Hai điều cần nhớ về hình này

Số nguồn cụ thể là số đo trên máy dựng cuốn sách này; máy khác ra con số khác. Thứ *không* đổi là hình dạng đường cong và vị trí chỗ gãy.

Đổi hướng của một nguồn nghĩa là đổi sang cặp đáp ứng xung khác. Đổi đột ngột thì nghe tiếng tách; phải chuyển trơn giữa hai cặp. Đó đúng là bài toán lập lịch tham số của chương 12.

10.8 Dự án thực hành — giai đoạn 3

10.8.1 Yêu cầu



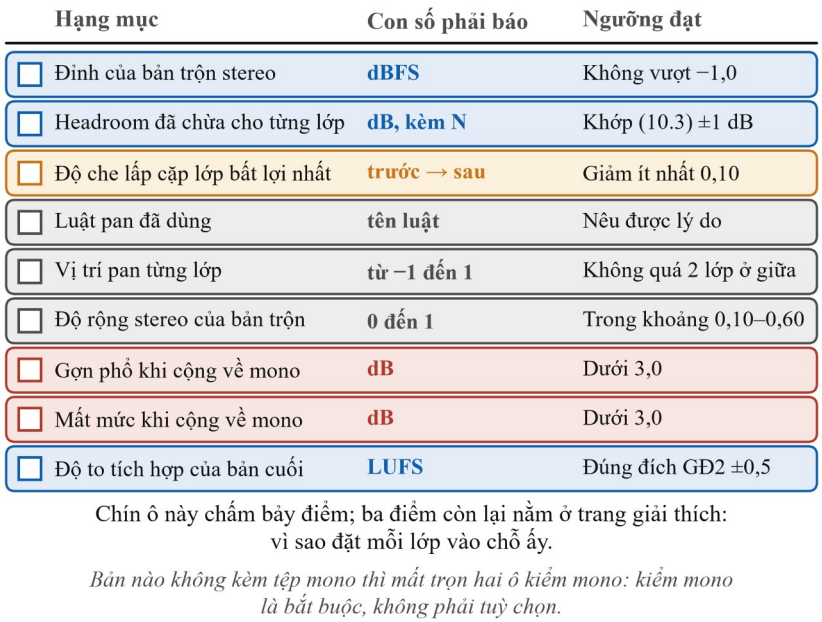
Hình 10.26 Năm bước của giai đoạn 3

Giai đoạn 3 nên làm xong trước chương 13, làm tiếp trên chính bản đã làm ở giai đoạn 2. Không cần tư liệu mới. Nếu bản giai đoạn 2 chỉ có một lớp thì tách nó ra thành nhiều lớp, hoặc thu thêm ba đoạn ngắn để có đủ bốn lớp.

Bảng 10.20 Kết quả cần có ở giai đoạn 3

Hạng mục	Yêu cầu cụ thể
Bốn tệp lớp rời	Ít nhất bốn lớp, mỗi lớp là một tệp WAV riêng, cùng tần số lấy mẫu và cùng độ dài
Bảng headroom	Số lớp N, lượng hạ đã dùng cho từng lớp, và đỉnh đo được sau khi cộng. So với công thức (10.3)
Bảng che lấp	Độ che lấp của cặp bất lợi nhất trước và sau khi xử lý, kèm cách đã dùng trong ba cách của mục 10.3.3
Bảng pan	Vị trí p của từng lớp, tên luật pan đã dùng, và lý do chọn luật ấy
Tệp trộn stereo	Bản trộn hoàn chỉnh, đỉnh không vượt −1,0 dBFS
Tệp trộn mono	Chính bản ấy cộng về mono, kèm gợn phổ và mất mức đo được
Một trang giải thích	Vì sao đặt mỗi lớp vào chỗ ấy. Đây là phần quan trọng nhất

10.8.2 Bảng tự kiểm trước khi coi là xong



Hình 10.27 Chín ô phải tự kiểm trước khi coi là xong

Chín ô trong hình đều máy kiểm được. Phần máy không kiểm được nằm ở trang giải thích: đặt lớp ở đâu không quan trọng bằng việc nêu được đã đo gì trước khi quyết định, và đã đánh đổi cái gì lấy cái gì.

Ba lỗi được dự đoán sẽ gặp nhiều nhất

Thiếu tệp mono. Không có nó thì hai ô cuối của bảng tự kiểm không kiểm được, và đó là hai ô kiểm đúng thứ dễ sai nhất.

Chừa headroom bằng cách hạ cả bản trộn *sau* khi cộng. Phần đã cắt đỉnh không lấy lại được — xem mục 10.2.5.

Pan cả sáu lớp ra hai bên cho "rộng". Độ rộng stereo vọt lên nhưng bản trộn rỗng ở giữa, và khi cộng mono thì gợn phổ vượt ngưỡng.

10.9 Thực hành, bài tập và tổng kết

10.9.1 Ba bài thực hành

Ba bài dưới đây làm ngay sau khi đọc xong chương, mỗi bài chừng mười phút, và cả ba đều phải *nghe* chứ không chỉ nhìn số. Bài thứ ba bắt buộc dùng tai nghe.

- **Bài 1 — headroom.** Lấy sáu tệp bất kỳ, chuẩn hoá cả sáu về −6 dBFS, rồi cộng ba lần: cộng thẳng, hạ 12 decibel mỗi lớp, hạ đúng theo công thức (10.3). Đo đỉnh của cả ba và nghe cả ba.
- **Bài 2 — pan và mono.** Pan hai lớp xung đột ra hai phía bằng cả ba luật pan. Với mỗi luật, đo mức ở giữa khi nghe stereo và khi cộng mono. Ba cặp con số phải khớp bảng ở mục 10.4.4.
- **Bài 3 — Haas và cái giá của nó.** Làm rộng một lớp bằng Haas ở 0,5 rồi 15 rồi 40 mili giây. Với từng lượng trễ: nghe ở stereo, nghe ở mono, và đo gợn phổ. Ghi lại lượng trễ nào cho kết quả tốt nhất trên *cả hai* cách nghe.

10.9.2 Bảy bài tập

Bảy bài dưới đây. Bài nào có đáp số thì đáp số in ngay sau đề — điểm nằm ở *cách làm* và ở phần giải thích, không nằm ở con số.

Bảng 10.21 Bảy bài tập của chương 10

#	Đề bài	Đáp số
1	Cần chừa headroom bao nhiêu cho 5 lớp độc lập? Cho 5 lớp đồng pha? Vì sao con số thứ hai gấp đôi con số thứ nhất?	−6,99 dB và −13,98 dB
2	Ở vị trí pan p = 0,5, tính cặp gain của cả ba luật, rồi tính mức nghe qua loa và mức khi cộng mono cho từng luật.	Giữ biên độ: −2,04 và 0,00 dB. Giữ công suất: 0,00 và +2,32 dB. Thoả hiệp: −1,03 và +1,15 dB
3	Một lớp làm rộng bằng Haas 3 ms, một lớp khác bằng Haas 25 ms. Rãnh đầu tiên khi cộng mono nằm ở đâu? Lớp nào nguy hơn và vì sao?	167 Hz và 20,0 Hz
4	Tính ITD cho nguồn ở 45 độ. Rồi đi ngược: một ITD đo được là 300 micro giây thì nguồn ở góc nào?	380,7 µs; và 34,7 độ
5	Tính ILD của nguồn ở 30 độ tại 300 Hz và tại 6 kHz. Giải thích chênh lệch bằng đại lượng ka.	1,88 dB và 9,89 dB
6	Đầu trẻ em có bán kính 7 centimet. Tính lại ITD lớn nhất, tần số giới hạn của ITD, và tần số ka bằng 1. Ba con số ấy đổi theo chiều nào so với người lớn?	524,7 µs; 1.225 Hz; 780 Hz
7	Tích chập theo khối 512 mẫu với đáp ứng xung HRIR 256 mẫu: so số phép nhân của cách thẳng và cách FFT cho *một* khối. Vì sao tỉ lệ ở đây nhỏ hơn hẳn con số 608 lần của mục 9.8?	Thẳng 131.072; FFT 31.744; nhanh 4,13 lần

10.9.3 Mở rộng cho ai muốn đi xa hơn

- **Bộ dò lỗi đảo cực.** Viết một hàm nhận vào danh sách lớp, tính hệ số tương quan của mọi cặp, và cảnh báo cặp nào cho giá trị âm sâu. Đây là công cụ mà mọi phòng thu đều nên có mà ít nơi có.
- **Pan phụ thuộc tần số.** Tách lớp thành hai dải bằng bộ lọc của mục 9.3, giữ dải trầm ở giữa và chỉ pan dải cao. Đo lại gợn phổ khi cộng mono và so với pan thường.

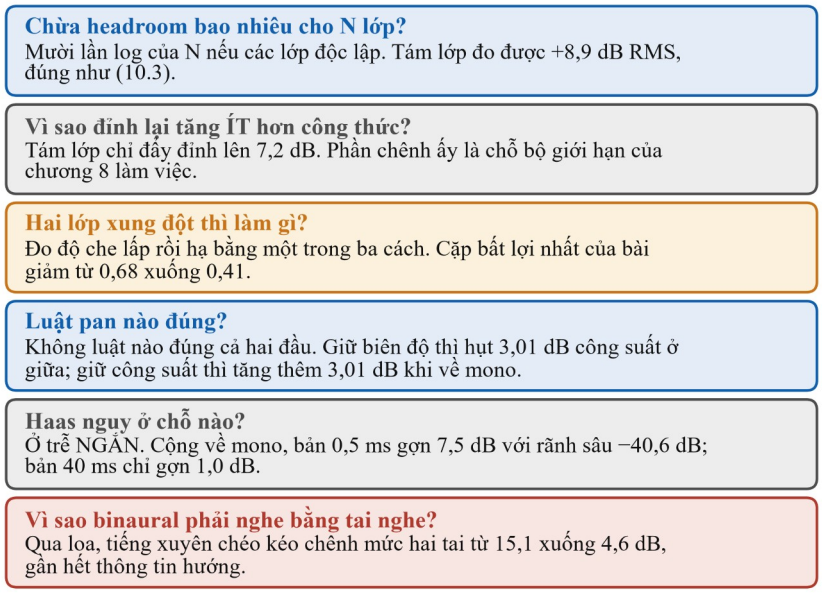
– **Bộ khử xuyên âm.** Cài đặt phép nghịch đảo của mô phỏng loa ở mục 10.7.4, để phát binaural qua hai loa. Đo lại chênh mức hai tai và xem khôi phục được bao nhiêu. Rồi thử dịch chỗ ngồi đi hai mươi centimet.
– **Nội suy giữa hai HRIR.** Dựng bảng đáp ứng xung cho các góc cách nhau 15 độ, rồi cho một nguồn chạy vòng quanh đầu. Thử ba cách chuyển: đổi đột ngột, làm mờ chéo, và nội suy hệ số. Nghe xem cách nào không có tiếng tách.

10.9.4 Câu hỏi tự trả lời

Bảy câu dưới đây không có đáp án in trong sách. Trả lời được cả bảy nghĩa là đã nắm chương này.

– Vì sao mức RMS của tám lớp độc lập tăng đúng như công thức, còn mức *đỉnh* thì tăng ít hơn? Chỗ chênh lệch ấy dùng vào việc gì?
– Hai lớp che lấp nhau 0,68 nhưng cả hai đều là lớp nền; hai lớp khác che lấp nhau 0,54 nhưng một trong hai là giọng hát. Xử lý cặp nào trước, và vì sao?
– Vì sao cắt giảm lớp này an toàn hơn nâng lớp kia, dù hai cách cho cùng một hiệu quả nghe được?
– Chứng minh bằng lời rằng không luật pan nào giữ phẳng được cả mức nghe qua loa lẫn mức khi cộng mono.
– Vì sao Haas trễ *ngắn* lại nguy hiểm hơn Haas trễ dài khi cộng về mono? Trả lời bằng khoảng cách giữa các rãnh.
– ITD lớn nhất chỉ 656 micro giây mà tai phân biệt được tới 10 micro giây. Con số ấy có ý nghĩa gì với cách viết phép trễ trong chương trình?
– Mô hình binaural đơn giản có ITD và ILD đúng, nhưng nguồn vẫn không phân biệt được trên với dưới. Thiếu cái gì, và vì sao thiếu nó thì sai?

10.9.5 Tổng kết — chương 10 để lại gì



Hình 10.28 Sáu câu phải trả lời được sau chương 10

Bảng 10.22 Những hàm mới của chương 10 và chỗ dùng lại

Hàm	Làm gì	Dùng lại ở
cong_lop	Cộng nhiều lớp, mỗi lớp một gain, tự đệm và tự lên kênh	Chương 12 — trộn động
headroom_can_db	Lượng phải hạ mỗi lớp cho N lớp	Dự án thực hành giai đoạn 3
he_so_tuong_quan	Hệ số tương quan Pearson giữa hai lớp	Dò lỗi đảo cực; đo độ rộng stereo
che_lap_pho	Độ che lấp phổ giữa hai lớp, trong [0, 1]	Chương 13 — đánh giá chất lượng tách nguồn
luat_pan · pan	Ba luật pan và phép áp chúng	Chương 11 — node panner của Web Audio
do_rong_stereo	Độ rộng stereo đo bằng tương quan hai kênh	Chương 14 — kiểm kết quả AI
gon_pho_khi_ve_mono_db	Gợn phổ và rãnh sâu nhất khi cộng mono	Dự án thực hành giai đoạn 3 và 4
haas	Làm rộng bằng trễ một kênh	Chương 11 — hiệu ứng thời gian thực
itd_giay · ild_db	Hai manh mối định vị	Chương 11 — đặt nguồn trong không gian trò chơi
binaural_don_gian · hrir_don_gian	Mô hình binaural đơn giản	Chương 11 và 12

Chương 10 khép lại bộ công cụ xử lý

Ba chương 8, 9 và 10 dựng xong toàn bộ bộ công cụ xử lý: khoảng động, tần số, thời gian và không gian.

Tất cả những thứ ấy tới giờ vẫn chạy theo lô — đọc cả tệp, xử lý, ghi ra tệp.

Chương 11 mở sang phần mới và đổi luật chơi: cùng những khối ấy nhưng phải chạy *thời gian thực*, trong trình duyệt, phản ứng theo sự kiện. Hình cuối của mục 10.7.5 là bản xem trước của bài toán ấy.

Tài liệu tham khảo của chương

Nguồn trực tiếp của các công thức trong bài

- Woodworth, R. S. — Experimental Psychology (1938). Nguồn của công thức xấp xỉ ITD ở mục 10.6.2. Công thức này đơn giản tới mức đáng ngờ nhưng vẫn là mô hình được dùng nhiều nhất, vì sai số của nó nhỏ hơn sai số giữa người này với người kia.
- Rayleigh, Lord (J. W. Strutt) — On Our Perception of Sound Direction, Philosophical Magazine (1907). Bài đặt nền cho thuyết song công ở mục 10.6.4, gồm cả lập luận về bóng đầu và về giới hạn pha.

- Haas, H. (1951) — Über den Einfluss eines Einfachechos auf die Hörsamkeit von Sprache. Bản dịch tiếng Anh đăng trên Journal of the Audio Engineering Society năm 1972. Nguồn của bảng năm mốc ở mục 10.5.2.
- Batteau, D. W. (1967) — The Role of the Pinna in Human Localization, Proceedings of the Royal Society B. Nguồn của mô hình phản xạ vành tai ở mục 10.7.3.
- Blauert, J. — Spatial Hearing: The Psychophysics of Human Sound Localization. Đây là sách tham khảo chuẩn cho toàn bộ mục 10.6 và 10.7, gồm cả nón lẫn lộn và độ nhạy hướng theo góc.

Sách

- Izhaki, R. — Mixing Audio: Concepts, Practices and Tools. Phần về headroom, về pan và về tương thích mono tương ứng đúng các mục 10.2 tới 10.5, viết cho người làm chứ không cho người đọc lý thuyết.
- Zölzer, U. (chủ biên) — DAFX: Digital Audio Effects. Chương về xử lý không gian có phần cài đặt HRTF và bộ khử xuyên âm.
- Begault, D. R. — 3-D Sound for Virtual Reality and Multimedia. Bản của NASA đọc miễn phí; đây là nguồn tốt nhất cho bốn giới hạn ở mục 10.7.4.
- Vorländer, M. — Auralization. Dành cho ai muốn nối phần vang của chương 9 với phần định vị của chương này thành một mô hình phòng đầy đủ.

Bộ dữ liệu và công cụ nên biết

- Bộ HRTF của CIPIC (Đại học California, Davis) và bộ *ari* của Viện Hàn lâm Áo. Cả hai đều tải miễn phí và đều có đo trên nhiều người — dùng để thấy tận mắt giới hạn thứ ba ở mục 10.7.4.
- Định dạng SOFA (Spatially Oriented Format for Acoustics), chuẩn AES69. Đây là định dạng chuẩn để lưu một bảng HRIR theo hướng; thư viện python-sofa đọc được nó.
- Node PannerNode và StereoPannerNode của Web Audio. Node thứ hai cài đúng luật giữ công suất ở mục 10.4.3; node thứ nhất có chế độ HRTF dựng sẵn. Cả hai là nội dung của chương 11.
- Bộ đo tương quan hai kênh (goniometer) có trong hầu hết phần mềm dựng nhạc. Nó vẽ trực tiếp đại lượng của mục 10.2.2, và là cách nhanh nhất để bắt lỗi đảo cực.

Đọc thêm theo hướng

- **Muốn đi sâu về trộn nhạc:** tìm đọc về nén song song và về trộn theo nhóm. Cả hai đều là cách quản lý headroom khi số lớp vượt quá mười sáu, tức chỗ mà bảng ở mục 10.2.3 dừng lại.
- **Muốn đi sâu về định vị:** tìm đọc về vai trò của chuyển động đầu. Quay đầu vài độ là ITD và ILD đổi theo, và chính sự thay đổi ấy giải quyết được nón lẫn lộn mà không cần tới manh mối phổ.

- **Muốn đi sâu về không gian nhiều loa:** tìm đọc về ambisonic và về VBAP. Ambisonic mô tả trường âm bằng các thành phần cầu chứ không bằng kênh, nên đổi số loa không phải trộn lại.
- **Muốn đi sâu về tích chập binaural:** tìm đọc về tích chập phân khối của Gardner — đúng công cụ để có trễ thấp ở đầu mà vẫn rẻ ở đuôi, tức là đi vòng qua chỗ đánh đổi của mục 10.7.5.

Một lưu ý về các con số trong bài

> **Cách đọc các con số**
>
> Các con số về headroom, về luật pan, về vị trí rãnh lọc lược, về ITD và ILD là kết quả tính chính xác từ công thức — chúng đúng với mọi tư liệu.
>
> Các con số về sáu lớp — độ che lấp, độ rộng stereo, gợn mono — là kết quả *đo* trên tư liệu dựng bằng mã, mô tả trong chú thích từng hình. Tư liệu khác sẽ cho con số khác nhưng chiều của quan hệ thì không đổi.
>
> Riêng số nguồn chạy nổi ở mục 10.7.5 là số đo thời gian thật trên máy dựng cuốn sách này; máy khác chắc chắn ra con số khác.
>
> Bảng ở mục 10.7.3 dựng từ một mô hình phản xạ vành tai đơn giản hoá, không phải số đo trên người thật.
>
> Mọi phép tính trong bài sinh ra từ mô-đun sigtools.py, đã qua 365 phép tự kiểm.