MỤC LỤC

Chương 14. Tích hợp ai vào sản phẩm 5
14.1 Vị trí của chương trong cuốn sách 6
14.2 Kết quả AI đến tay người dùng ở trạng thái nào 9
14.3 Đổi tần số lấy mẫu 16
14.4 Chuẩn hoá mức và đỉnh thật 23
14.5 Ghép đoạn và mối nối 28
14.6 Đường ống hậu kỳ — thứ tự và ngân sách 33
14.7 Đo và kiểm tự động 37
14.8 Dự án thực hành — giai đoạn 4 40
14.9 Thực hành, bài tập và tổng kết 43

DANH MỤC HÌNH VẼ

Hình 14.1 Bảy mục kỹ thuật của chương 14.....6
Hình 14.2 Bảy khối của đường ống hậu kỳ, và hai cách xếp sai hay gặp.....7
Hình 14.3 Một đoạn vừa nhận: dạng sóng ở trên, phổ trung bình ở dưới.....11
Hình 14.4 Ba đoạn vừa nhận, đo trên ba nhóm chỉ số.....12
Hình 14.5 Hai cách bỏ lệch một chiều, và vì sao chúng khác nhau.....13
Hình 14.6 Ba khối của cách hiểu, và một vòng lặp của cách làm.....18
Hình 14.7 Phổ của một hình sin 997 Hz sau ba cách đổi tần số.....19
Hình 14.8 Ba hệ quả của số cạnh, đo trên đường 44,1 lên 48 kHz.....20
Hình 14.9 Phần dải giữ nguyên vẹn trong phạm vi 0,1 decibel.....22
Hình 14.10 Cùng một tín hiệu đưa xuống 24 kHz theo hai cách.....23
Hình 14.11 Chuẩn hoá ba đoạn về −14 LUFS, và hậu quả với đỉnh thật.....24
Hình 14.12 Đỉnh thật nằm giữa hai mẫu, và hậu quả sau khi đổi tần số.....25
Hình 14.13 Ba mức trần, đo sau khi đã đổi lên 48 kHz.....27
Hình 14.14 Hai luật mờ chéo, mức ở giữa chỗ nối, và phép đo bốn ô.....29
Hình 14.15 Cùng một chỗ cắt, nối thẳng và nối bằng mờ chéo 20 mili giây.....31
Hình 14.16 Bước nhảy ở mối nối, trung vị trên 372 chỗ cắt.....31
Hình 14.17 Bốn cách xếp cùng năm khối, và kết quả.....34
Hình 14.18 Ba đường đổi tần số, đo thô và đo sau khi căn.....35
Hình 14.19 Hai ngân sách của cùng một đường ống.....36
Hình 14.20 Bộ kiểm bàn giao chạy trên đoạn 1.....38
Hình 14.21 Ba phép thử để đặt ra dung sai hồi quy.....39
Hình 14.22 Bốn kết quả cần có ở giai đoạn 4.....41
Hình 14.23 Các hạng mục tự đánh giá giai đoạn 4.....42
Hình 14.24 Chương 14 gói lại trong sáu con số.....44

DANH MỤC BẢNG BIỂU

Bảng 14.1 Bảy khối và câu hỏi riêng của mỗi khối....8
Bảng 14.2 Mười con số của chương và nơi đo chúng....9
Bảng 14.3 Năm khuyết tật cố tình đặt vào vật liệu....10
Bảng 14.4 Sáu chỉ số và điều mỗi chỉ số nói ra....11
Bảng 14.5 Hai cách bỏ lệch một chiều, đo trên đoạn 1....13
Bảng 14.6 Đáp ứng của bộ lọc thông cao 20 Hz tại các hài của vật liệu....14
Bảng 14.7 Dò vách trên sáu tần số gốc, sau khi đã nâng hết lên 48 kHz....16
Bảng 14.8 Tỉ số tối giản của sáu cặp tần số hay gặp....17
Bảng 14.9 Ba cách đổi tần số, đo trên cùng một hình sin....20
Bảng 14.10 Số cạnh và ba hệ quả của nó....21
Bảng 14.11 Dải giữ nguyên vẹn trong 0,1 decibel, theo số cạnh....22
Bảng 14.12 Bỏ bớt mẫu so với lọc rồi mới bỏ....23
Bảng 14.13 Chuẩn hoá ba đoạn về −14 LUFS....24
Bảng 14.14 Ba đoạn chuẩn hoá ĐỈNH về −0,1 dBFS rồi đo lại....25
Bảng 14.15 Ba cách đặt mức cho bản giao, đo trên đoạn 1....26
Bảng 14.16 Quét trần của bộ giới hạn....27
Bảng 14.17 Hai thứ tự, đo trên đoạn 1....27
Bảng 14.18 Hệ số tương quan giữa đuôi đoạn trước và đầu đoạn sau....28
Bảng 14.19 Bốn ô: hai loại nguồn nhân hai luật....29
Bảng 14.20 Bước nhảy ở mối nối theo độ dài mờ chéo....32
Bảng 14.21 Ghép ba đoạn bằng hai luật, mờ chéo 40 mili giây....33
Bảng 14.22 Quên cắt lặng thì bản ghép dài thêm bao nhiêu....33
Bảng 14.23 Bốn thứ tự, đo trên đoạn 1....34
Bảng 14.24 Đổi tần số nhiều chặng, ba cách đo....35
Bảng 14.25 Ngân sách độ trễ cho một đường ống chơi trực tiếp....37
Bảng 14.26 Thời gian máy, đo trên mỗi giây âm thanh....37
Bảng 14.27 Bảy tiêu chí và lý do của từng ngưỡng....38
Bảng 14.28 Bộ kiểm chạy trên ba đoạn, trước và sau đường ống....39
Bảng 14.29 Ba phép thử hồi quy trên cùng một đoạn....40
Bảng 14.30 Năm giai đoạn của dự án thực hành....41
Bảng 14.31 Bốn kết quả cần có và tiêu chuẩn của từng thứ....42

CHƯƠNG 14
TÍCH HỢP AI VÀO SẢN PHẨM

Mục tiêu của chương

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

- Đo được sáu chỉ số đầu tiên của một tệp vừa nhận từ mô hình, và nói được tệp ấy cần những khối xử lý nào.
- Phân biệt được hai cách bỏ lệch một chiều, và giải thích được vì sao một bộ lọc phẳng về độ lớn vẫn làm đổi đỉnh.
- Dò được vách phổ của một tệp và suy ra tần số mà nó thật sự được sinh ra.
- Rút gọn được tỉ số hữu tỉ của một phép đổi tần số, và tính được chi phí cùng độ trễ của bộ lọc đa pha.
- Cài được phép đổi tần số bằng bộ lọc đa pha, và đo được độ sạch của nó so với hai cách rẻ hơn.
- Chọn được số cạnh của bộ lọc có căn cứ, giữa ba thứ đánh đổi nhau.
- Giải thích được vì sao bỏ bớt mẫu mà không lọc là lỗi không sửa được, và đo được hậu quả của nó.
- Phân biệt được đỉnh mẫu với đỉnh thật, và chỉ ra được vì sao một tệp an toàn ở tần số này lại bị cắt đỉnh ở tần số khác.
- Chọn được mức giao hàng và trần của bộ giới hạn, dựa trên số đo chứ không dựa trên thói quen.
- Đo được hệ số tương quan giữa hai đoạn sắp ghép và chọn được luật mờ chéo đúng theo kết quả đo.
- Dựng được một phép đo có kiểm soát cho chỗ nối, và nhận ra được khi phép đo đang đo nhầm thứ khác.
- Xếp được năm khối của đường ống theo đúng thứ tự, và nói được mỗi thứ tự sai gây lỗi ở đâu.
- Lập được hai ngân sách — độ trễ và thời gian máy — cho một đường ống cụ thể.
- Viết được một bộ kiểm bàn giao có ngưỡng lập luận được, và đặt được dung sai hồi quy từ số đo.

Một chương không có khái niệm mới — và đó là chủ ý

Gần như mọi thứ dùng trong chương này đều đã học: định lý lấy mẫu ở chương 3, bộ lọc FIR ở chương 4, LUFS và đỉnh thật ở chương 7, bộ giới hạn ở chương 10, mờ chéo ở chương 12.

Cái mới nằm ở chỗ khác: *Thứ tự* các khối quyết định kết quả, và ba trong bốn thứ tự trông hợp lý đều cho ra một tệp không đạt. Mục 14.6 đo cả bốn.

Chương này cũng là chương đầu tiên của cuốn sách dạy cách viết một bộ kiểm: bảy tiêu chí, mỗi tiêu chí một con số đo được và một ngưỡng lập luận được. Đó là thứ người đọc sẽ dùng để tự kiểm dự án thực hành của mình.

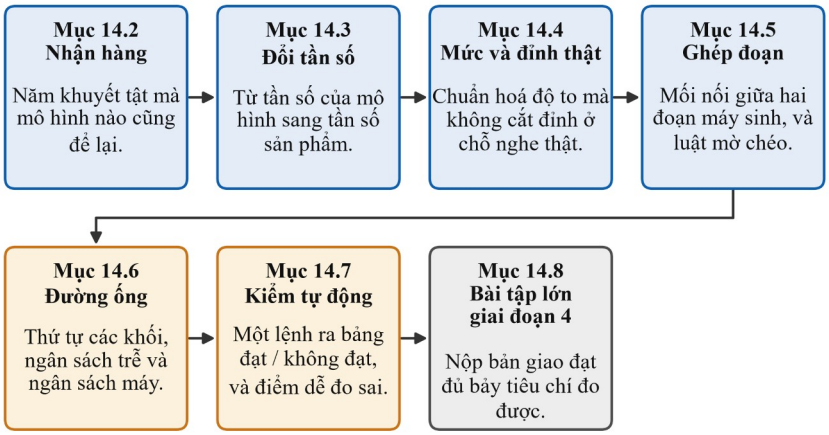
Mô-đun tính toán của cuốn sách có thêm ba mươi hai hàm cho chương này, và bộ tự kiểm 452 phép chạy bằng một lệnh, đạt 100 phần trăm.

14.1 Vị trí của chương trong cuốn sách

14.1.1 Chương 13 sinh ra âm thanh, chương này giao nó cho người khác

Chương 13 dừng lại đúng ở chỗ mô hình trả về một mảng số. Mảng ấy nghe được, nhưng nó chưa phải một sản phẩm. Nó ở sai tần số lấy mẫu, nó lệch một chiều, nó có vài trăm mili giây lặng ở hai đầu, và nếu ghép ba đoạn như thế lại thì mức của chúng chênh nhau gần bốn decibel.

Chương này làm đúng một việc: biến mảng số ấy thành một tệp có thể giao. Và cũng như mọi chương khác của cuốn sách, "có thể giao" ở đây không phải một cảm nhận mà là bảy con số đo được, mỗi con số có một ngưỡng.



Hình 14.1 Bảy mục kỹ thuật của chương 14

14.1.2 Mọi thứ cần dùng đều đã học rồi

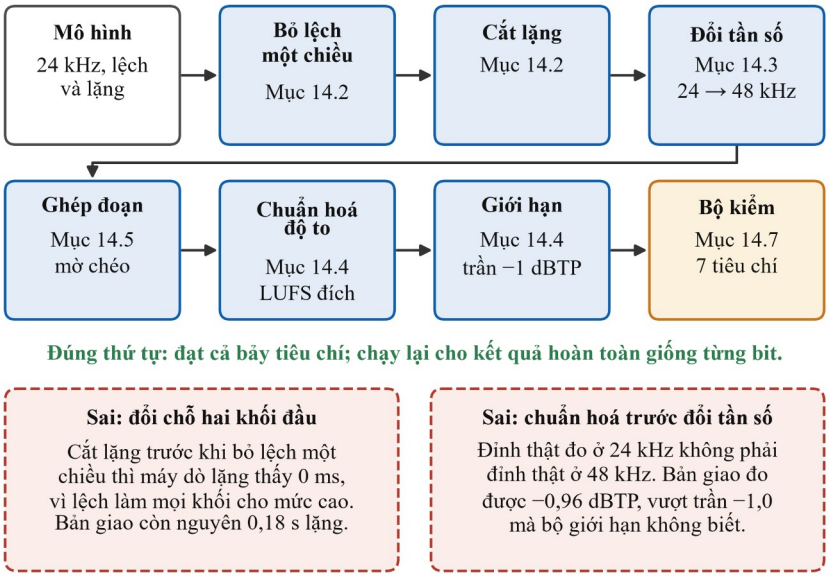
Điểm dễ chịu của chương này là nó gần như không có khái niệm mới. Nó lấy những thứ đã học ở mười ba chương trước rồi đặt chúng vào đúng thứ tự.

– **Lấy mẫu và chồng phổ ở mục 3.2.** Toàn bộ mục 14.3 là định lý lấy mẫu áp vào một việc rất cụ thể: đưa tín hiệu từ tần số này sang tần số khác mà không đẻ ra thứ không có trong bản gốc.

– **Bộ lọc FIR và tích chập ở mục 4.6.** Bộ lọc đổi tần số là một bộ lọc FIR bình thường; chỗ mới chỉ là cách xếp hệ số của nó để khỏi nhân với những số 0.
– **LUFS, đỉnh thật và bộ giới hạn ở mục 7.7 và 10.6.** Mục 14.4 dùng lại đúng ba thước đo ấy, nhưng lần này trên vật liệu do máy sinh — thứ có tỉ số đỉnh trên độ to cao hơn hẳn bản thu thường.
– **Mờ chéo và luật đẳng công suất ở mục 12.3.4.** Mục 14.5 chứng minh bằng số rằng đó đúng là luật phải dùng khi ghép hai đoạn do mô hình sinh — vì chúng không tương quan với nhau.

Cái mới thật sự của chương này không nằm ở công thức nào. Nó nằm ở chỗ: *Thứ tự* các phép xử lý quyết định kết quả, và ba trong bốn thứ tự hợp lý đều cho ra một tệp không đạt.

14.1.3 Đường ống hậu kỳ — bản đồ của cả chương



Hình 14.2 Bảy khối của đường ống hậu kỳ, và hai cách xếp sai hay gặp

Bảng 14.1 Bảy khối và câu hỏi riêng của mỗi khối

Khối	Việc của nó	Câu hỏi quyết định
Bỏ lệch một chiều	Đưa giá trị trung bình về 0	Trừ trung bình hay lọc thông cao — và vì sao khác nhau
Cắt lặng	Bỏ phần im ở hai đầu	Ngưỡng bao nhiêu, và chừa lại bao nhiêu
Đổi tần số	Từ tần số mô hình sang tần số sản phẩm	Bộ lọc dài bao nhiêu, và mất bao nhiêu dải
Ghép đoạn	Nối nhiều đoạn thành một bản	Luật mờ chéo nào, và dài bao nhiêu
Chuẩn hoá độ to	Đưa về mức giao hàng	LUFS đích bao nhiêu, và trần nào
Giới hạn	Không cho vượt trần	Trần đặt ở đâu để sau khi đổi tần số vẫn còn an toàn
Bộ kiểm	Nói đạt hay không đạt	Bao nhiêu tiêu chí, và ngưỡng lấy từ đâu

Một nguyên tắc chạy suốt chương

Mỗi khối trong đường ống đều thay đổi thứ mà khối sau *đo được*. Lệch một chiều làm máy dò lặng mù; chưa đổi tần số thì đỉnh thật đo ra chưa phải đỉnh thật của bản giao; chuẩn hoá xong mới giới hạn thì bộ giới hạn chỉ phải chạm vào vài gai lẻ.

Vì thế thứ tự không phải chuyện thẩm mỹ. Mục 14.6 đo bốn cách xếp và cho thấy hai trong số đó cho ra tệp không đạt — dù nghe thì không ai phân biệt nổi.

14.1.4 Mười con số sẽ đo được trong chương này

Bảng dưới đây là bản tóm tắt trước của cả chương. Mỗi dòng là một con số sẽ đo được, và mục nơi phép đo ấy nằm. Đọc lại bảng này sau khi học xong là cách nhanh nhất để kiểm tra đã nắm được gì.

Bảng 14.2 Mười con số của chương và nơi đo chúng

Con số	Giá trị đo được	Mục
Tỉ số tối giản của 44,1 sang 48 kHz	160 / 147	14.3.1
Rác của ba cách đổi tần số	−27,7 · −62,5 · −94,5 dB	14.3.4
Dải giữ nguyên vẹn khi đổi 48 xuống 24 kHz, c = 16	75,2 % của Nyquist	14.3.6
Hình sin 15 kHz khi bỏ bớt mẫu mà không lọc	gập xuống 9 kHz, gần nguyên mức	14.3.7
Tản độ to giữa ba đoạn vừa nhận	3,87 dB	14.4.1
Đỉnh thật của một đoạn đã chuẩn hoá đỉnh về −0,1 dBFS	+1,41 dBTP	14.4.2
Hệ số tương quan giữa hai đoạn liền nhau	−0,0011	14.5.1
Bước nhảy khi nối thẳng, so với mức bình thường	2,77 lần	14.5.3
Độ trễ lẻ do đi vòng qua 44,1 kHz	0,497 mẫu	14.6.2
Số tiêu chí phải đạt để gọi là một bản giao	7 / 7	14.7.1

14.2 Kết quả AI đến tay người dùng ở trạng thái nào

14.2.1 Vật liệu của chương và vì sao nó được dựng như thế

Cả chương làm việc trên ba đoạn nhạc do một mô hình giả lập sinh ra. Chúng không phải bản thu thật, và đó là chủ ý: mọi con số trong bài phải chạy lại được hoàn toàn như cũ trên máy của người đọc, không phụ thuộc vào một tệp âm thanh nào.

Nhưng vật liệu giả lập chỉ có ích nếu nó mang đúng những khuyết tật của vật liệu thật. Mô-đun sinh vật liệu của chương này cố tình đặt vào năm khuyết tật, và cả năm đều là thứ gặp hằng ngày khi lấy kết quả từ một mô hình sinh.

Bảng 14.3 Năm khuyết tật cố tình đặt vào vật liệu

Khuyết tật	Vì sao mô hình thật cũng có	Mục xử lý
Tần số lấy mẫu 24 kHz	Mô hình huấn luyện ở tần số thấp cho rẻ; sản phẩm cần 48 kHz	14.3
Lệch một chiều 0,004 – 0,012	Tầng dựng sóng không có ràng buộc nào buộc giá trị trung bình về 0	14.2.3
Mức trôi giữa các đoạn	Mỗi lần gọi mô hình là một lần độc lập, không có mốc mức chung	14.4
Lặng 0,18 – 0,32 giây ở hai đầu	Mô hình sinh đủ độ dài yêu cầu, phần thừa nó lấp bằng im lặng	14.2.4
Nền nhiễu khoảng −70 dB	Sai số lượng tử của tầng dựng sóng và của chính codec nơ-ron	14.7

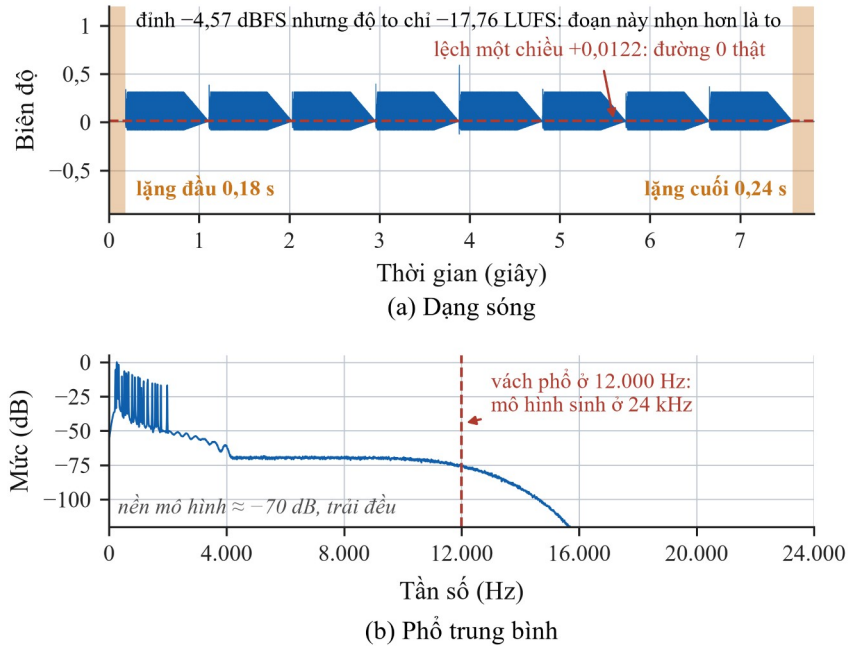
Liệt kê 14.1 Sinh một đoạn có đủ năm khuyết tật

```
def doan_ai(chi_so=0, dur=8.0, sr=24000):
    y = _giai_dieu(dur - 0.6, sr, bac, 220.0 * (1 + 0.03*chi_so))
    y = y + 0.004 * rng.standard_normal(y.size)   # nen mo hinh
    lang_dau  = np.zeros(int((0.18 + 0.05*chi_so) * sr))
    lang_cuoi = np.zeros(int((0.24 + 0.04*chi_so) * sr))
    y = np.concatenate([lang_dau, y, lang_cuoi])
    y = y * (0.58 - 0.09 * chi_so)                # muc troi
    y = y + (0.012 - 0.004 * chi_so)              # lech mot chieu
    return y
```

Ba đoạn khác nhau ở cả ba chỗ: giai điệu, mức, và độ lớn của từng khuyết tật. Đó là tình huống thật khi gọi một mô hình ba lần để lấy ba đoạn rồi ghép lại.

14.2.2 Sáu chỉ số đầu tiên phải đo trước khi làm bất cứ gì

Việc đầu tiên khi nhận một tệp từ mô hình không phải là nghe nó. Việc đầu tiên là đo sáu chỉ số. Sáu con số ấy nói ngay tệp này cần những khối xử lý nào, và tiết kiệm rất nhiều thời gian nghe đi nghe lại.



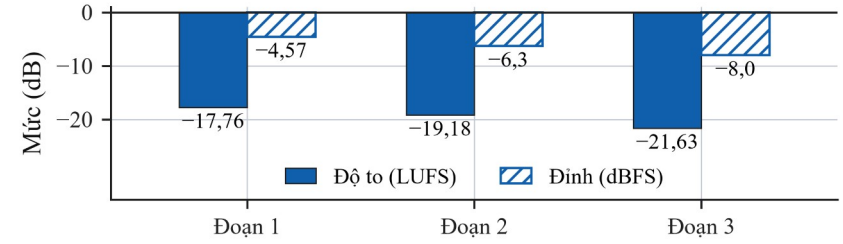
(a) Dạng sóng

(b) Phổ trung bình

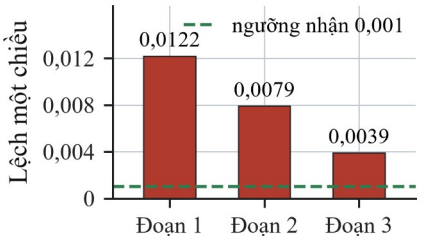
Hình 14.3 Một đoạn vừa nhận: dạng sóng ở trên, phổ trung bình ở dưới

Bảng 14.4 Sáu chỉ số và điều mỗi chỉ số nói ra

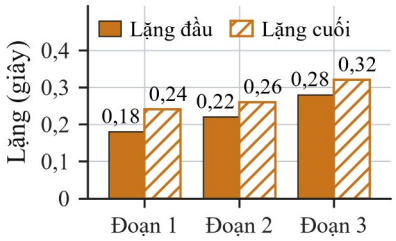
Chỉ số	Đo bằng gì	Nói ra điều gì
Đỉnh (dBFS)	peak_dbfs	Còn bao nhiêu chỗ trống trước khi bị cắt đỉnh
Độ to (LUFS)	do_to_tich_hop	Nghe to hay nhỏ so với bản khác
Lệch một chiều	lech_mot_chieu	Có phải dọn trước khi làm mọi phép đo khác không
Lặng hai đầu (giây)	lang_dau_cuoi	Ghép nhiều đoạn thì thừa ra bao nhiêu
Vách phổ (Hz)	phat_hien_tan_so_cat	Mô hình thật sự sinh ở tần số nào
Mẫu chạm trần (%)	ti_le_mau_cat	Đã bị cắt đỉnh từ trước khi đến tay người dùng chưa



(a) Độ to và đỉnh: tản độ to 3,87 dB



(b) Lệch quá ngưỡng ở cả ba đoạn



(c) Lặng thừa ở hai đầu

Hình 14.4 Ba đoạn vừa nhận, đo trên ba nhóm chỉ số

Ba đoạn có độ to lần lượt −17,76 · −19,18 · −21,63 LUFS. Chênh lệch giữa đoạn to nhất và đoạn nhỏ nhất là 3,87 decibel. Đó là mức chênh mà bất cứ ai cũng nghe ra ngay khi ba đoạn nối tiếp nhau: nó nghe như ba lần thu khác nhau chứ không phải một bản nhạc.

Đáng chú ý hơn là quan hệ giữa đỉnh và độ to. Đoạn 1 có đỉnh −4,57 dBFS mà độ to chỉ −17,76 LUFS: chênh nhau hơn mười ba decibel. Vật liệu do máy sinh thường nhọn hơn bản thu thật, và mục 14.4 sẽ cho thấy đó chính là thứ làm cho phép chuẩn hoá độ to trở nên nguy hiểm.

14.2.3 Lệch một chiều: hai cách dọn, hai kết quả khác nhau

Lệch một chiều là phần giá trị trung bình khác 0 của tín hiệu. Nó không nghe thấy — tai người không đáp ứng ở 0 Hz — nhưng nó làm sai lệch gần như mọi phép đo và mọi khối xử lý đặt sau nó.

$$d = \frac{1}{N}\sum_{n=0}^{N-1} x[n], \qquad x_{\text{sạch}}[n] = x[n] - d \tag{14.1}$$

Ở đây d là giá trị trung bình, N là số mẫu, và x sạch là tín hiệu sau khi trừ đi phần lệch. Đoạn 1 của vật liệu có d bằng 0,0122 — nghĩa là toàn bộ dạng sóng bị đẩy lên trên đường 0 một khoảng bằng 1,2 phần trăm toàn thang.

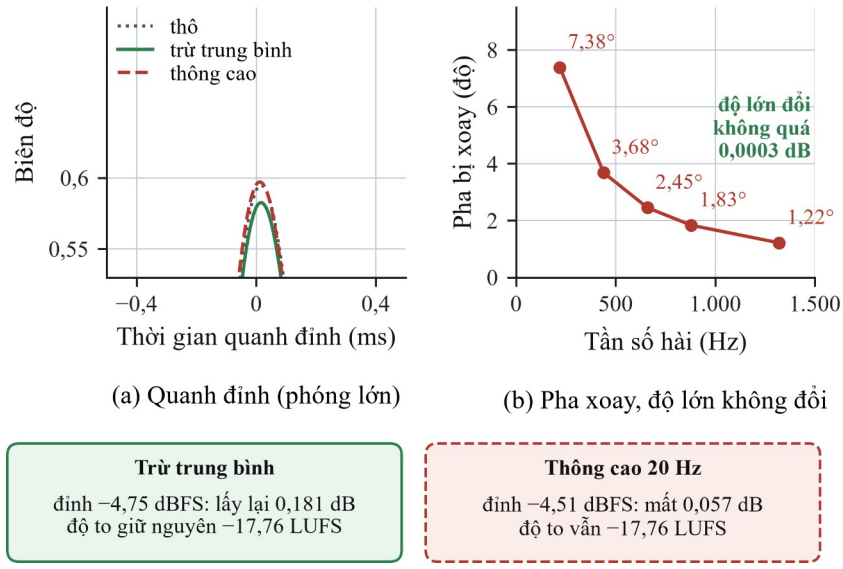
Ba thứ sai vì lệch một chiều

Máy dò lặng đọc sai. Một khối "im lặng" có lệch 0,0122 vẫn cho mức hiệu dụng −38 dB, cao hơn ngưỡng lặng rất nhiều — nên máy dò báo 0 mili giây lặng ở cả hai đầu, trong khi thật ra có 180 và 240 mili giây.

Đỉnh đo sai. Phần lệch cộng thẳng vào biên độ, nên tệp mất chỗ trống ở một phía và thừa ở phía kia.

Mọi khối sau phải tải thêm một thành phần một chiều. Bộ nén và bộ giới hạn phản ứng với nó; bộ lọc thông thấp giữ nguyên nó; và khi ghép hai đoạn có lệch khác nhau thì chỗ nối có một bậc thang không bao giờ mất đi.

Có hai cách dọn, và chúng *không* cho cùng một kết quả.



Hình 14.5 Hai cách bỏ lệch một chiều, và vì sao chúng khác nhau

Bảng 14.5 Hai cách bỏ lệch một chiều, đo trên đoạn 1

	Trừ trung bình	Thông cao 20 Hz
Lệch còn lại	0,0000	0,0000
Đỉnh sau khi dọn	−4,75 dBFS	−4,51 dBFS
So với trước (−4,57)	lấy lại 0,181 dB chỗ trống	*Mất* 0,057 dB chỗ trống
Độ to sau khi dọn	−17,76 LUFS	−17,76 LUFS
Xử lý được lệch *trôi* theo thời gian	Không	Có
Có làm xoay pha không	Không	Có

Chỗ đáng ngạc nhiên là dòng thứ ba. Bộ lọc thông cao 20 Hz *làm tăng* đỉnh của tín hiệu, dù nó không hề khuếch đại gì. Lần đầu gặp hiện tượng này rất dễ đổ cho quá độ ở đầu tệp — nhưng đỉnh mới không nằm ở đầu tệp.

Câu trả lời nằm ở đáp ứng của bộ lọc. Đo tại năm hài đầu của vật liệu thì độ lớn gần như không đổi, còn pha thì đổi thật.

Bảng 14.6 Đáp ứng của bộ lọc thông cao 20 Hz tại các hài của vật liệu

Tần số (Hz)	Độ lớn bị đổi (dB)	Pha bị xoay (độ)
220	−0,0003	7,38
440	−0,0000	3,68
660	−0,0000	2,45
880	−0,0000	1,83
1.320	−0,0000	1,22

Đỉnh phụ thuộc pha, độ to thì không

Bộ lọc giữ nguyên độ lớn của mọi hài tới 0,0003 decibel, nhưng xoay pha hài bậc nhất đi 7,38 độ và hài bậc hai đi 3,68 độ. Các hài vì thế cộng lại theo một cách hơi khác, và tổng của chúng có một đỉnh cao hơn 0,057 decibel.

Độ to tích hợp không đổi một chút nào — nó đo năng lượng, mà năng lượng không phụ thuộc pha. Đây là minh hoạ sạch nhất trong cả cuốn sách cho câu: *Đỉnh* và *độ to* là hai đại lượng khác nhau về bản chất, không phải hai cách nói về cùng một thứ.

Hệ quả thực hành: mọi khối chỉ đổi pha — bộ lọc pha tuyến tính thì không, nhưng bộ lọc bậc hai thông thường thì có — đều có thể làm tệp vượt trần dù không khuếch đại gì. Vì thế đo đỉnh phải đo ở *cuối* đường ống, không đo ở giữa.

Cách dùng trong bài là làm cả hai: trừ trung bình để bỏ phần lệch cố định, rồi thông cao 20 Hz để bắt phần lệch trôi theo thời gian. Đó là chế độ mặc định của hàm trong mô-đun.

Liệt kê 14.2 Bỏ lệch một chiều — ba chế độ

```
def bo_lech_mot_chieu(x, sr=SR, f_cat=20.0, cach="ca_hai"):
    if cach in ("tru_trung_binh", "ca_hai"):
        x = x - float(np.mean(x))
    if cach in ("thong_cao", "ca_hai"):
        b, a = biquad_he_so("thong_cao", f_cat, 0.7071, 0.0, sr)
        x = ap_biquad(x, b, a)
    return x
```

14.2.4 Lặng hai đầu — và một bài học về thứ tự

Mô hình sinh gần như luôn để lại lặng ở hai đầu. Với một đoạn thì không sao; với mười đoạn ghép lại thì đó là vài giây im lặng nằm rải rác giữa bản nhạc.

Cách dò rất đơn giản: chia tín hiệu thành các khối 20 mili giây, tính mức hiệu dụng của từng khối, và tìm khối đầu tiên cùng khối cuối cùng có mức cao hơn đỉnh 60 decibel.

Liệt kê 14.3 Dò lặng ở hai đầu

```
def lang_dau_cuoi(x, sr=SR, nguong_db=-60.0, cua_ms=20.0):
    n   = max(int(cua_ms / 1000.0 * sr), 1)
    so  = x.size // n
    khoi = x[:so*n].reshape(so, n)
    muc  = 20*np.log10(np.maximum(np.sqrt((khoi**2).mean(1)), 1e-12))
    keu  = np.where(muc > muc.max() + nguong_db)[0]
    return keu[0]*n/sr, (so - 1 - keu[-1])*n/sr
```

Chạy hàm ấy trên đoạn thô cho ra 0,00 và 0,00 giây — nghĩa là "không có lặng". Chạy đúng hàm ấy trên đoạn đã bỏ lệch một chiều cho ra 0,18 và 0,24 giây, đúng bằng thứ đã đặt vào.

Một lỗi trượt có thể che một lỗi khác

Khối "im lặng" của đoạn thô có giá trị 0,0122 ở mọi mẫu, nên mức hiệu dụng của nó là −38,3 dB. Đỉnh của đoạn là −4,57 dB, nên ngưỡng lặng nằm ở −64,6 dB. Khối im lặng cao hơn ngưỡng ấy tới 26 decibel, và máy dò kết luận nó *không* lặng.

Đây là lý do vì sao bỏ lệch một chiều phải là khối *đầu tiên* của đường ống, không phải một khối nào đó ở giữa. Mọi phép đo đặt trước nó đều có nguy cơ đọc ra một con số trông hợp lý mà sai.

Bộ kiểm ở mục 14.7 tái hiện đúng điểm dễ sai này: bản thô *đạt* tiêu chí "lặng đầu" chỉ vì nó *trượt* tiêu chí "lệch một chiều".

Khi cắt, phải chừa lại một khoảng — bài dùng 20 mili giây. Cắt sát đúng mẫu đầu tiên vượt ngưỡng sẽ chặt mất phần lên của nốt đầu, và tạo ra đúng cái bậc thang mà mục 14.5 sẽ bàn.

14.2.5 Vách phổ: mô hình thật sự sinh ở tần số nào

Một tệp WAV ghi 48.000 Hz ở phần đầu tệp không có nghĩa là nội dung của nó thật sự có thông tin tới 24.000 Hz. Nếu mô hình sinh ở 24 kHz rồi ai đó nâng lên 48 kHz, tệp vẫn ghi 48.000 nhưng phần trên 12.000 Hz chỉ còn nền nhiễu.

Dấu vết ấy nhìn thấy được trên phổ trung bình: có một *vách*, một chỗ mức tụt hẳn xuống rồi không bao giờ lên lại. Điều quan trọng là phải dò vách chứ không dò theo một ngưỡng tuyệt đối — vì mức nền của mỗi tệp mỗi khác.

$$f_{\text{vách}} = \min f \text{ saocho } \bar{S}(g) \le \bar{S}_{\max} - \Delta \text{ vớimọi } g \ge f, \qquad \Delta = 25 \text{ dB} \tag{14.2}$$

Công thức đọc là: vách nằm ở tần số nhỏ nhất mà từ đó trở lên, *mọi* tần số đều thấp hơn đỉnh phổ ít nhất 25 decibel. Cách cài đặt gọn nhất là lấy bao trên ngược bằng cộng dồn cực đại từ phải sang trái, rồi tìm chỗ bao ấy cắt ngưỡng.

Liệt kê 14.4 Dò vách phổ bằng bao trên ngược

```
def phat_hien_tan_so_cat(x, sr=SR, sut_db=25.0, n_fft=8192, muot=9):
    f, db = pho_trung_binh_db(x, sr, n_fft)
    db = np.convolve(db, np.ones(muot)/muot, mode="same")
    bao = np.maximum.accumulate(db[::-1])[::-1]   # bao tren NGUOC
    duoi = np.where(bao <= db.max() - sut_db)[0]
    return float(f[duoi[0]]) if duoi.size else float(f[-1])
```

Bảng 14.7 Dò vách trên sáu tần số gốc, sau khi đã nâng hết lên 48 kHz

Tần số mô hình sinh	Nyquist của nó	Vách dò được	Sai số
8.000 Hz	4.000 Hz	3.885 Hz	−2,9 %
16.000 Hz	8.000 Hz	8.314 Hz	+3,9 %
22.050 Hz	11.025 Hz	11.713 Hz	+6,2 %
24.000 Hz	12.000 Hz	12.861 Hz	+7,2 %
32.000 Hz	16.000 Hz	17.555 Hz	+9,7 %
48.000 Hz	24.000 Hz	24.000 Hz	0,0 %

Sai số vài phần trăm là do dải chuyển tiếp của bộ lọc nâng tần: vách thật không phải một bức tường thẳng đứng mà là một con dốc. Nhưng độ chính xác ấy quá đủ để trả lời câu hỏi cần trả lời — mô hình này sinh ở 16 kHz hay 24 kHz — và để phát hiện ra rằng tệp 48 kHz vừa nhận thật ra chỉ có nội dung tới 12 kHz.

> **Dùng vách phổ để làm gì trong thực tế**
>
> Biết mô hình sinh ở tần số nào thì biết được đổi tần số bằng tỉ số nào — mục 14.3 cho thấy tỉ số 2/1 rẻ hơn tỉ số 160/147 tới tám mươi lần.
>
> Biết được rằng nâng tần *không* thêm thông tin. Một tệp 24 kHz nâng lên 48 kHz vẫn là nội dung 24 kHz, chỉ tốn gấp đôi dung lượng. Đừng quảng cáo nó là 48 kHz.
>
> Biết được nguồn của một tệp lạ. Vách ở 11 kHz thường là dấu vết của 22,05 kHz; vách ở 15,5 kHz là dấu vết quen thuộc của MP3 128 kbit mỗi giây — thứ chương 15 sẽ dùng để tìm vi phạm bản quyền.

14.3 Đổi tần số lấy mẫu

14.3.1 Tỉ số hữu tỉ: hai con số quyết định mọi thứ còn lại

Đổi từ tần số này sang tần số khác không phải một phép nội suy tuỳ ý. Nó là một phép toán có cấu trúc, và cấu trúc ấy do một phân số quyết định.

$$g = \gcd(f_{\text{vào}}, f_{\text{ra}}), \qquad L = \frac{f_{\text{ra}}}{g}, \qquad M = \frac{f_{\text{vào}}}{g} \tag{14.3}$$

Rút gọn tỉ số giữa tần số ra và tần số vào cho ra hai số nguyên L và M. Toàn bộ phép đổi là: chèn thêm $L-1$ số 0 vào giữa mỗi cặp mẫu, lọc thông thấp, rồi giữ lại mỗi M mẫu một mẫu. Hai con số ấy quyết định cả chi phí lẫn độ khó của bộ lọc.

Bảng 14.8 Tỉ số tối giản của sáu cặp tần số hay gặp

Từ	Sang	L	M	Nhận xét
44.100	48.000	160	147	Đắt — hai số đều lớn
48.000	44.100	147	160	Đắt như trên, ngược chiều
24.000	48.000	2	1	Rẻ nhất — chỉ là gấp đôi
22.050	48.000	320	147	Đắt nhất trong bảng
16.000	44.100	441	160	Rất đắt
32.000	48.000	3	2	Rẻ

Một quyết định kiến trúc rất rẻ

Nếu được chọn tần số của mô hình, hãy chọn một tần số có tỉ số đơn giản với tần số sản phẩm. Mô hình sinh ở 24 kHz đưa lên 48 kHz tốn L = 2; mô hình sinh ở 22,05 kHz đưa lên cùng chỗ ấy tốn L = 320.

Chênh lệch ấy không đổi chất lượng — cả hai đều làm được sạch — nhưng nó đổi bộ nhớ cần cho bộ lọc và độ dài của nó. Đây là loại quyết định mất năm giây khi thiết kế và tốn rất nhiều về sau nếu chọn sai.

14.3.2 Ba khối nối tiếp — và vì sao không ai làm đúng như thế

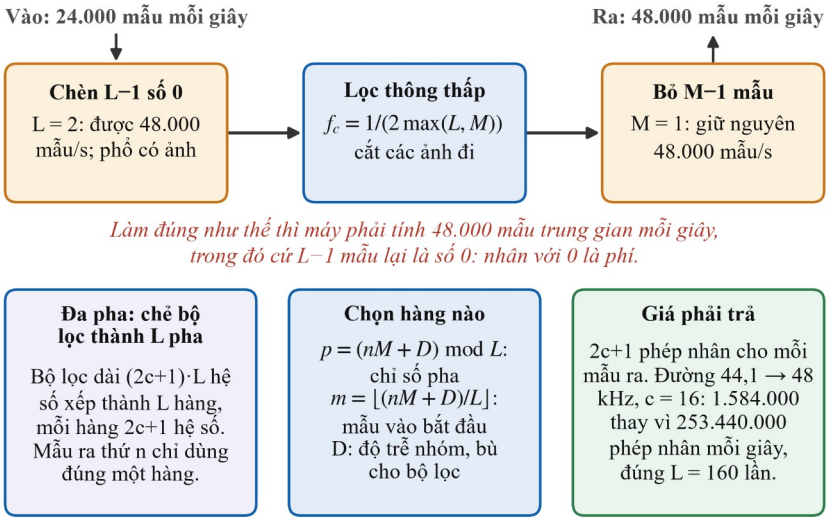
Cách hiểu chuẩn của phép đổi tần số gồm ba khối.

$$x_L[m] = x[m/L] \ \text{khi}\ L \mid m\,; \qquad x_L[m] = 0 \ \text{cácchỉsốcònlại} \tag{14.4}$$

Khối thứ nhất chèn số 0: tín hiệu trung gian có L lần số mẫu, và cứ L−1 mẫu lại là một số 0. Phổ của nó chứa L bản sao của phổ gốc — gọi là các *ảnh*.

$$f_c = \frac{1}{2\max(L, M)} \ \left(\text{theotầnsốtrunggian}\ L f_{\text{vào}}\right) \tag{14.5}$$

Khối thứ hai là bộ lọc thông thấp cắt hết các ảnh ấy. Tần số cắt phải nhỏ hơn cả nửa Nyquist của đầu vào lẫn nửa Nyquist của đầu ra, nên nó lấy theo số lớn hơn trong hai số L và M. Khối thứ ba giữ lại mỗi M mẫu một mẫu.



Hình 14.6 Ba khối của cách hiểu, và một vòng lặp của cách làm

Làm đúng ba khối ấy thì phải tính 48.000 mẫu trung gian mỗi giây, mà một nửa trong số đó là số 0. Nhân một hệ số bộ lọc với số 0 là phí, và khi L bằng 160 thì 159 trên 160 phép nhân là phí.

14.3.3 Bộ lọc đa pha: cùng một kết quả, ít phép nhân hơn L lần

Ý tưởng của đa pha đơn giản đến mức dễ bị coi thường: nếu biết trước mẫu nào là số 0 thì đừng tính chúng. Chẻ bộ lọc dài (2c+1)·L hệ số thành L hàng, mỗi hàng 2c+1 hệ số. Mỗi mẫu ra chỉ dùng đúng một hàng.

$$y[n] = \sum_{k=0}^{K-1} g_p[k]\, x[m-k], \qquad p = (nM + D) \bmod L, \qquad m = \left\lfloor \frac{nM + D}{L} \right\rfloor \tag{14.6}$$

Ở đây p là chỉ số hàng — gọi là *pha* — và m là mẫu vào nơi phép tích chập bắt đầu. D là độ trễ nhóm của bộ lọc, cộng vào để tín hiệu ra căn đúng với tín hiệu vào.

Liệt kê 14.5 Toàn bộ phép đổi tần số bằng đa pha

```
def doi_sr_da_pha(x, L, M, so_canh=16, beta=8.6):
    h   = bo_loc_doi_sr(L, M, so_canh, beta)
    K   = h.size // L
    G   = h.reshape(K, L).T              # (L, K) — L pha
    tre = (h.size - 1) // 2
    n   = np.arange(int(np.floor(x.size * L / M)))
    t   = n * M + tre
    p, m = t % L, t // L                 # pha nao, bat dau o dau
    idx = (m[:, None] - np.arange(K)[None, :]) + dem
    return (xp[idx] * G[p]).sum(axis=1)
```

Bộ lọc dùng cửa sổ Kaiser nhân với hàm sinc — cách chuẩn mực để dựng một bộ lọc thông thấp pha tuyến tính có độ triệt điều chỉnh được bằng một tham số.

Liệt kê 14.6 Bộ lọc thông thấp của phép đổi tần số

```
def bo_loc_doi_sr(L, M, so_canh=16, beta=8.6):
    K  = 2 * int(so_canh) + 1
    N  = K * L
    n  = np.arange(N) - (N - 1) / 2.0
    fc = 0.5 / max(L, M)
    h  = 2.0*fc*np.sinc(2.0*fc*n) * np.kaiser(N, float(beta))
    return h * L                          # bu lai muc cho phep chen 0
```

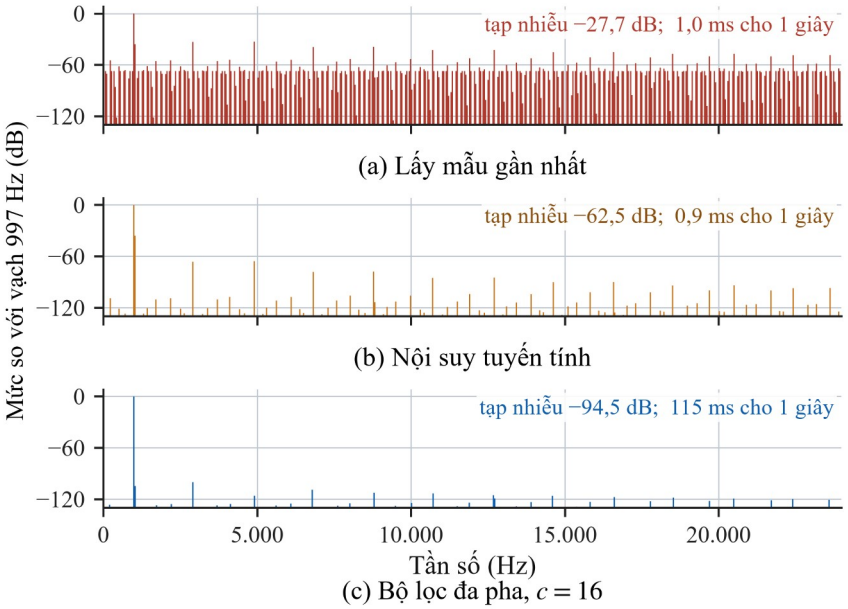
$$C_{\text{nhân}} = (2c+1) f_{\text{ra}}, \qquad \tau_{\text{lọc}} = \frac{(2c+1)L - 1}{2 L f_{\text{vào}}} \tag{14.7}$$

Chi phí là 2c+1 phép nhân cho mỗi mẫu ra, không phụ thuộc L. Với đường 44,1 lên 48 kHz và c bằng 16 thì đó là 1.584.000 phép nhân mỗi giây, thay vì 253.440.000 nếu làm theo đúng ba khối — đúng 160 lần, bằng L.

14.3.4 Ba cách đổi, và khoảng cách 67 decibel giữa chúng

Để thấy vì sao phải làm phức tạp như thế, hãy so với hai cách rẻ hơn. Cách thứ nhất: với mỗi mẫu ra, lấy mẫu vào gần nhất. Cách thứ hai: nội suy tuyến tính giữa hai mẫu vào gần nhất.

Phép đo dùng một hình sin 997 Hz sạch đưa từ 44,1 lên 48 kHz. Mọi vạch phổ khác 997 Hz đều do phép đổi đẻ ra, nên tổng của chúng là thước đo trực tiếp cho mức tạp nhiễu.



Hình 14.7 Phổ của một hình sin 997 Hz sau ba cách đổi tần số

Bảng 14.9 Ba cách đổi tần số, đo trên cùng một hình sin

Cách	Rác còn lại	Thời gian cho 1 giây	Dùng khi nào
Lấy mẫu gần nhất	−27,7 dB	1,0 ms	Không bao giờ
Nội suy tuyến tính	−62,5 dB	0,9 ms	Chỉ khi đổi tỉ số liên tục, ví dụ đổi cao độ trực tiếp
Bộ lọc đa pha c = 16	−94,5 dB	115 ms	Mặc định cho mọi việc hậu kỳ

Khoảng cách giữa cách rẻ nhất và cách đắt nhất là 67 decibel. Cái giá là 115 mili giây cho mỗi giây âm thanh — nghĩa là dựng một bản ba phút mất khoảng hai mươi giây. Đó là cái giá không đáng bàn cãi với công việc hậu kỳ.

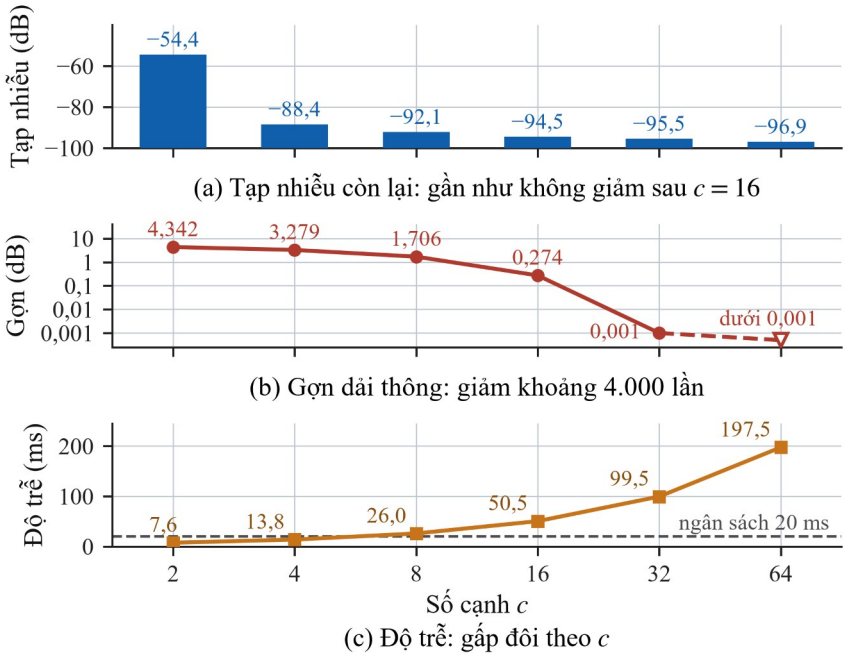
Rác của phép đổi tần số không sửa được về sau

Vạch tạp mà cách lấy mẫu gần nhất sinh ra nằm ở những tần số không có quan hệ hài với tín hiệu gốc. Không có bộ lọc nào bóc chúng ra được, vì chúng nằm lẫn trong dải nghe.

Vì thế đây là một trong vài chỗ của cả cuốn sách mà lời khuyên là tuyệt đối: dùng bộ lọc đa pha, hoặc dùng một thư viện đã cài nó. Đừng tự viết nội suy tuyến tính cho việc đổi tần số.

14.3.5 Chọn số cạnh: ba thứ đánh đổi nhau

Tham số c — số cạnh của bộ lọc — quyết định bộ lọc dài bao nhiêu. Nó kéo theo ba hệ quả, và ba hệ quả ấy không cùng chiều.



(a) Tạp nhiễu còn lại: gần như không giảm sau $c = 16$

(b) Gợn dải thông: giảm khoảng 4.000 lần

(c) Độ trễ: gấp đôi theo c

Hình 14.8 Ba hệ quả của số cạnh, đo trên đường 44,1 lên 48 kHz

Bảng 14.10 Số cạnh và ba hệ quả của nó

c	Tap mỗi mẫu ra	Rác còn lại	Gợn dải thông	Độ trễ
2	5	−54,4 dB	4,342 dB	7,6 ms
4	9	−88,4 dB	3,279 dB	13,8 ms
8	17	−92,1 dB	1,706 dB	26,0 ms
16	33	−94,5 dB	0,274 dB	50,5 ms
32	65	−95,5 dB	0,001 dB	99,5 ms
64	129	−96,9 dB	dưới 0,001 dB	197,5 ms

Đọc bảng theo cột thì thấy ba câu chuyện khác nhau. Độ triệt ảnh gần như đứng yên sau c bằng 16 — thêm bốn lần số phép nhân chỉ mua thêm 2,4 decibel. Gợn dải thông thì ngược lại: nó giảm bốn nghìn lần từ c bằng 2 đến c bằng 32. Còn độ trễ thì cứ gấp đôi theo c, không có chỗ nào bão hoà.

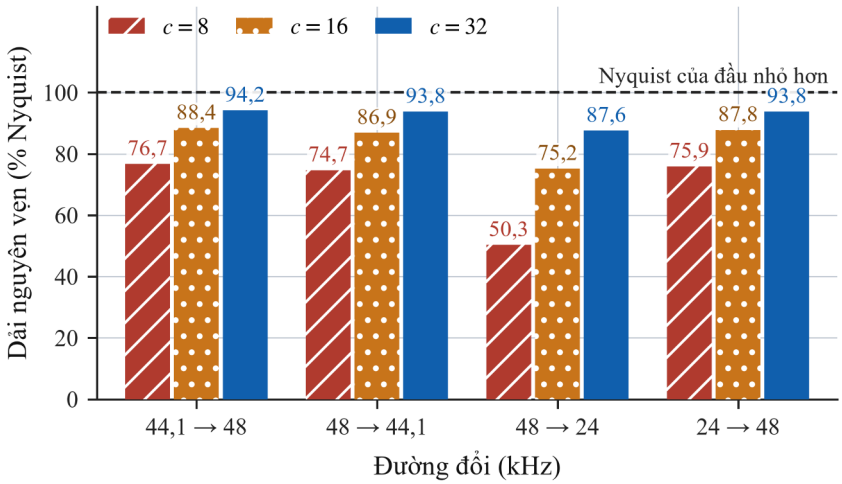
> **Vì sao c = 16 là mặc định hợp lý**
>
> Độ triệt −94,5 decibel đã dưới mức nghe được trong mọi hoàn cảnh thực tế, và thêm nữa không mua được gì.
>
> Gợn dải thông 0,274 decibel nằm dưới ngưỡng nghe ra chênh lệch mức của tai người — khoảng 0,5 đến 1 decibel với tín hiệu phức tạp.
>
> Độ trễ 50 mili giây là nhiều cho việc chơi trực tiếp, nhưng hậu kỳ thì không quan tâm — và với đường 24 lên 48 kHz thì bộ lọc ngắn hơn rất nhiều, chỉ 0,34 mili giây, vì L chỉ bằng 2.

14.3.6 Dải thật sự giữ được — và nó không bằng Nyquist

Có một hiểu lầm dai dẳng: đưa tín hiệu từ 48 xuống 24 kHz thì giữ được mọi thứ dưới 12.000 Hz. Điều đó chỉ đúng nếu bộ lọc là một bức tường thẳng đứng, mà không bộ lọc nào như thế.

$$B_{\text{giữ}} = \frac{\min(f_{\text{vào}}, f_{\text{ra}})}{2}(1 - \rho) \tag{14.8}$$

Hệ số rho là phần dải bị dải chuyển tiếp ăn mất. Nó phụ thuộc vào c và vào tỉ số L trên M, và cách duy nhất để biết nó là đo.



Hình 14.9 Phần dải giữ nguyên vẹn trong phạm vi 0,1 decibel

Bảng 14.11 Dải giữ nguyên vẹn trong 0,1 decibel, theo số cạnh

Đường đổi	Nyquist	c = 8	c = 16	c = 32
44,1 → 48 kHz	22.050 Hz	16.904 Hz (76,7 %)	19.488 Hz (88,4 %)	20.780 Hz (94,2 %)
48 → 44,1 kHz	22.050 Hz	16.473 Hz (74,7 %)	19.165 Hz (86,9 %)	20.672 Hz (93,8 %)
48 → 24 kHz	12.000 Hz	6.042 Hz (50,3 %)	9.023 Hz (75,2 %)	10.512 Hz (87,6 %)
24 → 48 kHz	12.000 Hz	9.114 Hz (75,9 %)	10.534 Hz (87,8 %)	11.262 Hz (93,8 %)

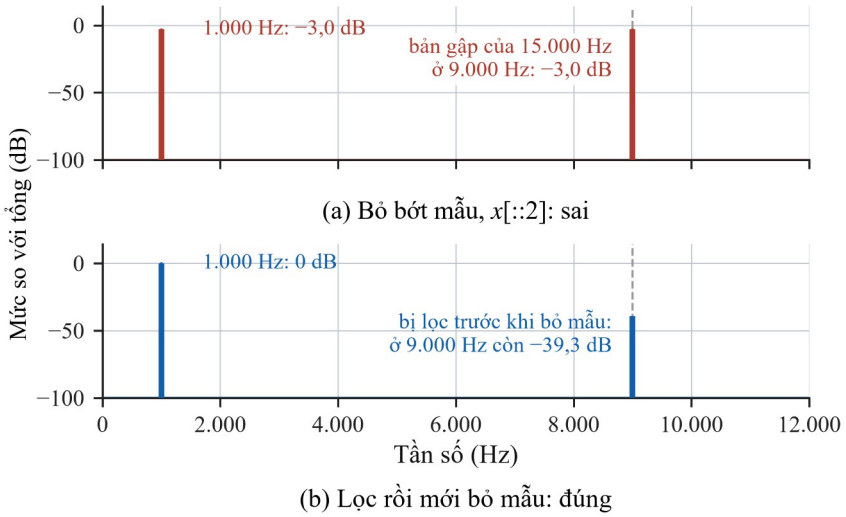
Hàng thứ ba là hàng đáng chú ý. Đi *xuống* 24 kHz với c bằng 8 thì chỉ giữ nguyên vẹn được một nửa dải — mọi thứ trên 6 kHz đã bắt đầu bị suy giảm. Đó là vì bộ lọc phải dốc hơn nhiều khi đi xuống: nó vừa phải giữ dải dưới 12 kHz vừa phải chặn đủ sạch trên 12 kHz để không có chồng phổ.

> **Quy tắc thực hành cho việc đi xuống**
>
> Khi đi xuống một tần số thấp hơn, dùng c lớn hơn bình thường — c bằng 32 thay vì 16. Cái giá là gấp đôi số phép nhân, và phần thưởng là mười hai phần trăm dải.
>
> Và nhớ rằng phép đi xuống *không* đảo ngược được. Phần dải mất đi không lấy lại được bằng cách nâng lên trở lại.

14.3.7 Bỏ sót việc lọc trước khi bỏ bớt mẫu: lỗi tốn kém nhất

Nếu đường đổi là 48 xuống 24 kHz thì M bằng 2 và L bằng 1, và có một cám dỗ rất lớn: chỉ cần lấy mỗi hai mẫu một mẫu, x[::2]. Nó chạy tức thì và trông có vẻ đúng.

Phép đo dùng một tín hiệu gồm hai hình sin: 1.000 Hz — nằm an toàn trong dải — và 15.000 Hz, nằm trên Nyquist mới là 12.000 Hz.



Hình 14.10 Cùng một tín hiệu đưa xuống 24 kHz theo hai cách

Bảng 14.12 Bỏ bớt mẫu so với lọc rồi mới bỏ

	Bỏ bớt mẫu (x[::2])	Lọc rồi mới bỏ
Thành phần ngoài 1.000 Hz	−3,0 dB	−39,3 dB
Hình sin 15.000 Hz đi đâu	Gập xuống 9.000 Hz, gần nguyên mức	Bị cắt trước khi bỏ mẫu
Có nghe ra không	Có — nó là một nốt lạ không có trong bản gốc	Không
Sửa được sau khi lưu tệp không	Không	—

Hình sin 15.000 Hz không biến mất. Nó gập xuống 24.000 trừ 15.000 bằng 9.000 Hz, và xuất hiện ở đó với gần như nguyên vẹn độ lớn. Trong bản nhạc thật, thứ gập xuống không phải một hình sin mà là toàn bộ phần cao của phổ, và nó tạo ra một lớp âm kim loại không có quan hệ hài với nhạc.

Đây là lỗi không sửa được

Sau khi đã lưu tệp 24 kHz, phần gập xuống nằm lẫn trong dải nghe và không phân biệt được với nội dung thật. Không bộ lọc nào, không mô hình nào lấy lại được bản gốc.

Mọi lỗi khác trong chương này đều sửa được bằng cách chạy lại đường ống từ bản gốc. Lỗi này thì không — trừ khi vẫn còn giữ bản 48 kHz.

Quy tắc: giữ bản gốc ở tần số cao nhất hiện có, và chỉ đổi xuống ở bước cuối cùng khi xuất bản giao.

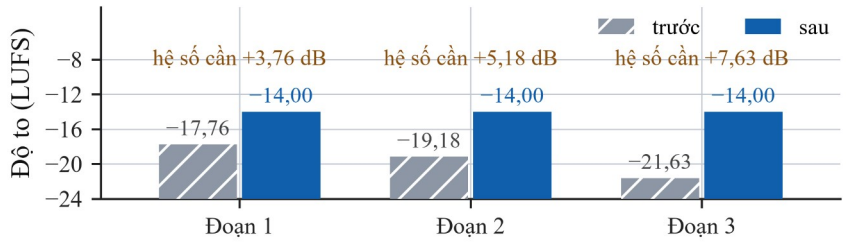
14.4 Chuẩn hoá mức và đỉnh thật

14.4.1 Đưa ba đoạn về cùng một độ to

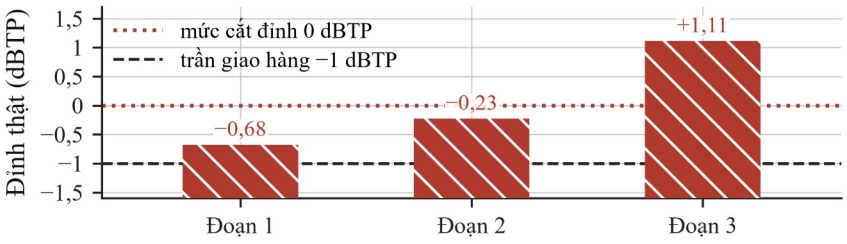
Ba đoạn của vật liệu chênh nhau 3,87 decibel về độ to. Việc đưa chúng về cùng một mức là phép nhân với một hệ số, và hệ số ấy tính thẳng từ hiệu giữa mức đích và mức đo được.

$$G = L_{\text{đích}} - L_{\text{đo}}, \qquad G_{\text{thật}} = \min\left(G,\ T_{\text{trần}} - P_{\text{tp}}\right) \tag{14.9}$$

Nhưng không được nhân đúng bằng G. Nếu nhân đúng bằng G mà đỉnh thật vượt trần thì phải hạ thêm — và khi ấy độ to không đạt đích nữa. Đó là lý do công thức có phần lấy giá trị nhỏ hơn.



(a) Độ to trước và sau khi chuẩn hoá về −14 LUFS



(b) Đỉnh thật sau chuẩn hoá: cả ba đoạn vượt trần

Hình 14.11 Chuẩn hoá ba đoạn về −14 LUFS, và hậu quả với đỉnh thật

Bảng 14.13 Chuẩn hoá ba đoạn về −14 LUFS

Đoạn	Trước	Hệ số cần	Sau	Đỉnh thật sau
1	−17,76 LUFS	+3,76 dB	−14,00 LUFS	−0,68 dBTP
2	−19,18 LUFS	+5,18 dB	−14,00 LUFS	−0,23 dBTP
3	−21,63 LUFS	+7,63 dB	−14,00 LUFS	+1,11 dBTP

Tản độ to giảm từ 3,87 decibel xuống đúng 0. Nhưng cột cuối cho thấy cái giá: cả ba đoạn đều vượt trần giao hàng −1 dBTP, và đoạn 3 — chính đoạn phải lên nhiều nhất — vượt cả mức cắt đỉnh 0 dBTP.

Chuẩn hoá độ to không tự lo phần đỉnh

Độ to đo năng lượng trung bình có trọng số theo tai; đỉnh đo giá trị tức thời lớn nhất. Hai đại lượng ấy chỉ liên hệ lỏng lẻo qua tỉ số đỉnh trên độ to, mà tỉ số ấy khác nhau tuỳ vật liệu.

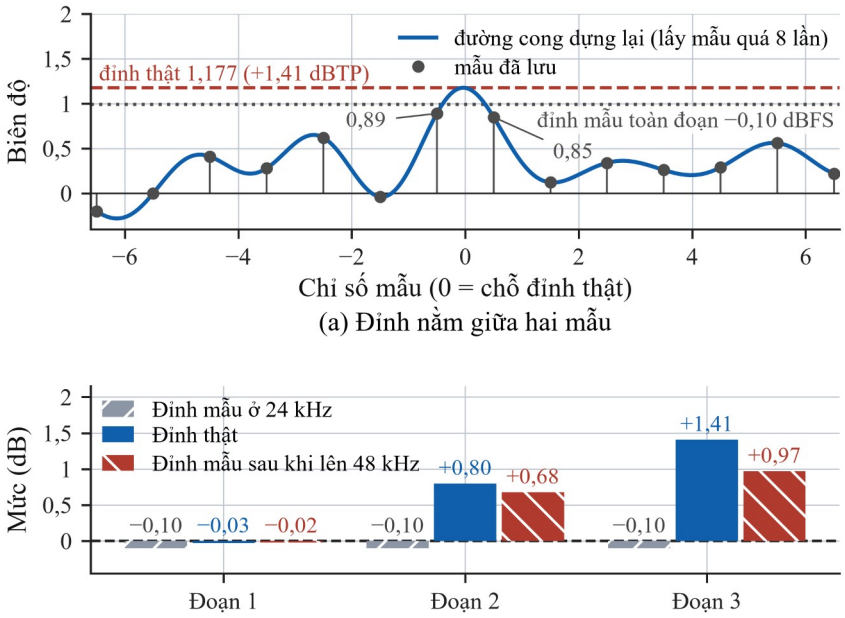
Vật liệu do mô hình sinh thường có tỉ số đỉnh trên độ to cao hơn bản thu thường — ba đoạn ở đây lần lượt là 13,3 · 13,8 · 15,1 decibel. Nghĩa là cùng một mức LUFS đích, chúng chạm trần sớm hơn một bản nhạc đã qua phòng thu.

Vì thế mọi phép chuẩn hoá độ to phải đi kèm một trần, và mục 14.4.4 đo xem trần ấy nên đặt ở đâu.

14.4.2 Đỉnh thật: thứ nằm giữa hai mẫu

Đỉnh mẫu chỉ nhìn những mẫu đã có. Nhưng bộ chuyển số sang tương tự không phát ra các mẫu ấy — nó dựng lại một đường cong đi qua chúng, và đường cong ấy có thể vọt cao hơn ở giữa hai mẫu. Chuẩn BS.1770 gọi giá trị lớn nhất của đường cong ấy là *đỉnh thật*, và yêu cầu đo bằng cách lấy mẫu quá ít nhất bốn lần.

$$P_{\text{tp}} = 20\log_{10}\max_{n}|x_{\uparrow 4}[n]|, \qquad \text{PLR} = P_{\text{tp}} - L_{\text{đo}} \tag{14.10}$$



(a) Đỉnh nằm giữa hai mẫu

(b) Chuẩn hoá đỉnh về −0,1 dBFS rồi đo lại

Hình 14.12 Đỉnh thật nằm giữa hai mẫu, và hậu quả sau khi đổi tần số

Hình (a) là một chỗ thật trong vật liệu: hai mẫu liền nhau ở 0,89 và 0,85, còn sóng thật đi giữa chúng vọt tới 1,177 — tức +1,41 dBTP, trong khi đỉnh mẫu chỉ là −0,10 dBFS.

Bảng 14.14 Ba đoạn chuẩn hoá ĐỈNH về −0,1 dBFS rồi đo lại

Đoạn	Đỉnh mẫu ở 24 kHz	Đỉnh thật	Chênh	Đỉnh mẫu sau khi lên 48 kHz
1	−0,10 dBFS	−0,03 dBTP	+0,07 dB	−0,02 dBFS
2	−0,10 dBFS	+0,80 dBTP	+0,90 dB	+0,68 dBFS
3	−0,10 dBFS	+1,41 dBTP	+1,51 dB	+0,97 dBFS

Cột cuối là cột quan trọng nhất của cả mục

Trước khi đổi tần số, đỉnh *mẫu* của cả ba đoạn đều là −0,10 dBFS — nghĩa là không có mẫu nào chạm trần, và mọi phần mềm đều báo tệp hoàn toàn an toàn.

Sau khi đổi lên 48 kHz, chính đỉnh *mẫu* của đoạn 2 và đoạn 3 vượt 0 dBFS. Phép đổi tần số đã lấy mẫu tại những vị trí mới, và một số vị trí ấy rơi đúng vào chỗ sóng đang vọt.

Nói cách khác: đỉnh thật *không* phải một khái niệm lý thuyết. Nó là đỉnh mẫu của tệp sau khi ai đó đổi tần số, đổi định dạng, hay cho qua một codec có mất mát — và mọi việc ấy đều xảy ra trên đường tệp đi tới người nghe.

14.4.3 Ba cách giao mức, và cái nào vượt trần

Có ba cách đặt mức cho bản giao, và chúng trả lời ba câu hỏi khác nhau.

Bảng 14.15 Ba cách đặt mức cho bản giao, đo trên đoạn 1

Cách	Độ to đạt được	Đỉnh thật	Sau khi lên 48 kHz	Kết luận
Chuẩn hoá *đỉnh* về −0,1 dBFS	−13,35 LUFS	−0,03 dBTP	đỉnh −0,02 dBFS · 0,000 % mẫu cắt	Đạt
Chuẩn hoá LUFS về −14	−14,00 LUFS	−0,68 dBTP	đỉnh −0,67 dBFS · 0,000 % mẫu cắt	Đạt
Chuẩn hoá LUFS về −9	−9,00 LUFS	+4,32 dBTP	đỉnh +4,33 dBFS · 0,201 % mẫu cắt	Vỡ

Cách thứ ba là cách mà một người muốn bản nhạc "nghe to bằng bản trên dịch vụ nhạc" hay chọn. Nó làm 0,201 phần trăm số mẫu chạm trần — khoảng 377 mẫu trong mỗi giây — và mỗi mẫu ấy là một chỗ méo.

Mức giao hàng nên lấy từ đâu

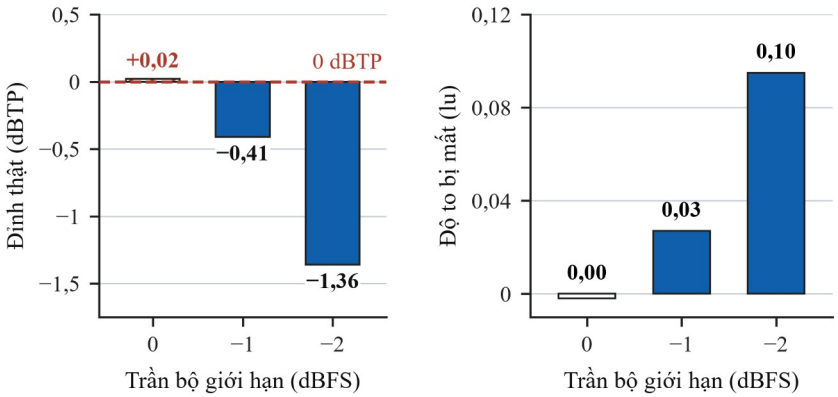
Các nền tảng phát nhạc đều chuẩn hoá độ to khi phát, nên đẩy bản nhạc to hơn mức đích của họ không làm nó nghe to hơn — chỉ làm nó nghe méo hơn sau khi bị hạ xuống.

Mức đích thường gặp cho nhạc trên nền tảng là khoảng −14 LUFS; cho phát thanh và truyền hình châu Âu là −23 LUFS theo EBU R 128. Bài này dùng −14 LUFS với trần −1 dBTP.

Điều cần nhớ không phải con số cụ thể — nó đổi theo nơi giao — mà là nguyên tắc: mức đích và trần là *hai* tham số, và tham số thứ hai có quyền phủ quyết tham số thứ nhất.

14.4.4 Trần của bộ giới hạn: 0 dBFS vẫn chưa đủ

Bộ giới hạn có nhìn trước ở mục 10.6 đảm bảo không mẫu nào vượt trần. Câu hỏi còn lại là đặt trần ở đâu. Cám dỗ là đặt ở 0 dBFS để không mất chút độ to nào.



(a) Đỉnh thật (b) Cái giá về độ to

Hình 14.13 Ba mức trần, đo sau khi đã đổi lên 48 kHz

Bảng 14.16 Quét trần của bộ giới hạn

Trần đặt ở	Độ to bị mất	Đỉnh thật sau khi lên 48 kHz	Kết luận
0 dBFS	0,00 *lu*	+0,02 dBTP	Còn vượt
−1 dBFS	0,03 *lu*	−0,41 dBTP	An toàn
−2 dBFS	0,10 *lu*	−1,36 dBTP	An toàn, hơi phí

Hạ trần từ 0 xuống −1 dBFS làm bản nhạc nhỏ đi 0,03 *lu*. Không ai nghe ra 0,03 *lu* — ngưỡng nghe ra chênh lệch độ to là khoảng 0,5 đến 1 *lu*. Đổi lại, đỉnh thật sau khi đổi tần số đi từ +0,02 về −0,41 dBTP.

Ba mức trần nên nhớ

−1 dBTP là trần giao hàng chuẩn cho hầu hết nền tảng, và là con số bài này dùng.

−2 dBTP nếu biết trước bản nhạc còn phải qua một codec có mất mát — MP3 hay AAC làm đỉnh thật tăng thêm vài phần mười decibel nữa.

0 dBFS chỉ dùng cho bản làm việc nội bộ, không bao giờ cho bản giao.

14.4.5 Thứ tự giữa chuẩn hoá và đổi tần số

Mục này đo một câu hỏi cụ thể: chuẩn hoá trước rồi đổi tần số, hay đổi tần số trước rồi chuẩn hoá.

Bảng 14.17 Hai thứ tự, đo trên đoạn 1

	Chuẩn hoá trước	Đổi tần số trước
Độ to đạt được	−14,05 LUFS	−14,00 LUFS
Đỉnh	−0,67 dBFS	−0,62 dBFS
Đỉnh thật	−0,63 dBTP	−0,59 dBTP
Sai so với đích	0,047 *lu*	0,000 *lu*

Chênh lệch nhỏ, nhưng nó chỉ một chiều: chuẩn hoá trước luôn trượt đích, vì phép đổi tần số sau đó thay đổi cả độ to lẫn đỉnh một chút. Đổi tần số trước rồi

chuẩn hoá thì đạt đích đúng bằng 0, vì không còn khối nào phía sau làm xê dịch nữa.

> **Một quy tắc chung cho cả đường ống**
>
> Mọi phép *đo* và mọi phép *đặt mức* phải nằm ở *cuối* đường ống, sau khi mọi khối làm thay đổi tín hiệu đã chạy xong.
>
> Lý do đơn giản: thứ chúng ta hứa với người nhận là con số của *tệp giao*, không phải con số của một bản trung gian nào đó.
>
> Mục 14.6 đo thêm hai cách xếp sai nữa, và cả hai đều sai vì cùng một lý do này.

14.5 Ghép đoạn và mối nối

14.5.1 Hai đoạn do mô hình sinh có tương quan với nhau không

Mục 12.3.4 đã dựng hai luật mờ chéo và nói rằng luật nào đúng phụ thuộc vào việc hai nguồn có tương quan hay không. Ở đây có một tình huống cụ thể để trả lời câu hỏi ấy bằng số.

Bảng 14.18 Hệ số tương quan giữa đuôi đoạn trước và đầu đoạn sau

Cặp đoạn	Hệ số tương quan	Kết luận
Đuôi đoạn 1 với đầu đoạn 2	−0,0011	Không tương quan — hai lần gọi mô hình độc lập
Hai nửa của cùng một đoạn	+0,0027	Cũng không tương quan — vì là hai chỗ khác nhau trong bản nhạc

Con số −0,0011 gần 0 đến mức không cần bàn: hai đoạn do mô hình sinh trong hai lần gọi khác nhau không có quan hệ pha nào với nhau. Đó chính là ô đầu tiên trong bảng bốn ô của mục 12.3.4, và nó chỉ thẳng ra luật phải dùng.

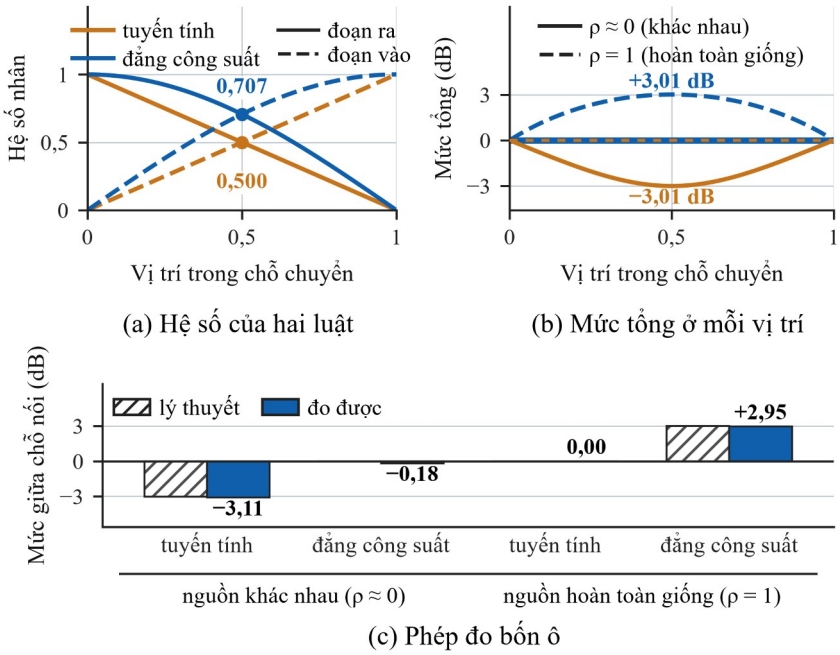
14.5.2 Hai luật mờ chéo, và phép đo khớp lý thuyết cả bốn ô

$$f_{\text{ra}}(t) = \cos\frac{\pi t}{2}, \qquad f_{\text{vào}}(t) = \sin\frac{\pi t}{2}, \qquad t \in [0,\ 1] \tag{14.11}$$

Luật tuyến tính dùng hai đường thẳng; luật đẳng công suất dùng cosin và sin. Ở giữa chỗ chuyển, luật tuyến tính cho hai hệ số bằng 0,500, còn luật đẳng công suất cho 0,707.

$$P(t) = f_{\text{ra}}^2 + f_{\text{vào}}^2 + 2\rho f_{\text{ra}} f_{\text{vào}} \tag{14.12}$$

Công thức tính mức tổng ở mỗi vị trí trong chỗ chuyển. Số hạng thứ ba chứa hệ số tương quan rho, và chính nó quyết định luật nào đúng. Khi rho bằng 0 thì luật đẳng công suất giữ mức không đổi; khi rho bằng 1 thì luật tuyến tính mới là luật giữ mức.



(a) Hệ số của hai luật (b) Mức tổng ở mỗi vị trí

(c) Phép đo bốn ô

Hình 14.14 Hai luật mờ chéo, mức ở giữa chỗ nối, và phép đo bốn ô

Bảng 14.19 Bốn ô: hai loại nguồn nhân hai luật

Nguồn	Luật	Lý thuyết	Đo được	Sai
Khác nhau (rho ≈ 0)	Tuyến tính	−3,01 dB	−3,11 dB	0,10 dB
Khác nhau (rho ≈ 0)	Đẳng công suất	0,00 dB	−0,18 dB	0,18 dB
Giống hệt (rho = 1)	Tuyến tính	0,00 dB	+0,00 dB	0,00 dB
Giống hệt (rho = 1)	Đẳng công suất	+3,01 dB	+2,95 dB	0,06 dB

Ba dòng này giải thích cả một loại lỗi

Ghép hai đoạn *không* tương quan bằng luật tuyến tính cho một chỗ hõm 3 decibel ở giữa chỗ nối. Ba decibel là rất nhiều: nó nghe như bản nhạc bị hụt hơi một nhịp.

Ghép hai đoạn *hoàn toàn giống* bằng luật đẳng công suất cho một chỗ *vồng* lên 3 decibel — cũng sai như thế, chỉ ngược chiều. Đây là tình huống khi mờ chéo hai bản thu cùng một nguồn, ví dụ hai lần lấy cùng một đoạn nền.

Không có luật nào đúng cho mọi trường hợp. Cách chọn đúng là *đo* hệ số tương quan giữa hai đoạn sắp ghép, rồi chọn theo. Với kết quả do mô hình sinh thì câu trả lời gần như luôn là đẳng công suất.

Liệt kê 14.7 Nối hai đoạn bằng mờ chéo, chọn được luật

```
def noi_mo_cheo(a, b, ms=40.0, sr=SR, luat="cong_suat"):
    n = min(int(ms / 1000.0 * sr), a.size, b.size)
    if n <= 1:
        return np.concatenate([a, b])
    t = np.arange(n) / float(n - 1)
    if luat == "tuyen_tinh":
        fo, fi = 1.0 - t, t
```

```
    else:
        fo, fi = np.cos(np.pi*t/2.0), np.sin(np.pi*t/2.0)
    giua = a[-n:] * fo + b[:n] * fi
    return np.concatenate([a[:-n], giua, b[n:]])
```

14.5.3 Đo tiếng tách ở mối nối — và cách đo cho đúng

Câu hỏi tiếp theo là mờ chéo phải dài bao nhiêu. Để trả lời được, cần một thước đo cho "chỗ nối lộ đến mức nào", và thước đo ấy khó hơn vẻ ngoài của nó.

Hai phép đo sai đã thử trước khi ra phép đo đúng

Đo năng lượng cao tần quanh chỗ nối, như đã làm ở mục 6.6. Phép đo ấy là một *tỉ lệ* trên tổng năng lượng, nên nó phụ thuộc vào nội dung nhạc ngay chỗ ấy nhiều hơn phụ thuộc vào chỗ nối. Kết quả không đơn điệu theo độ dài mờ chéo.

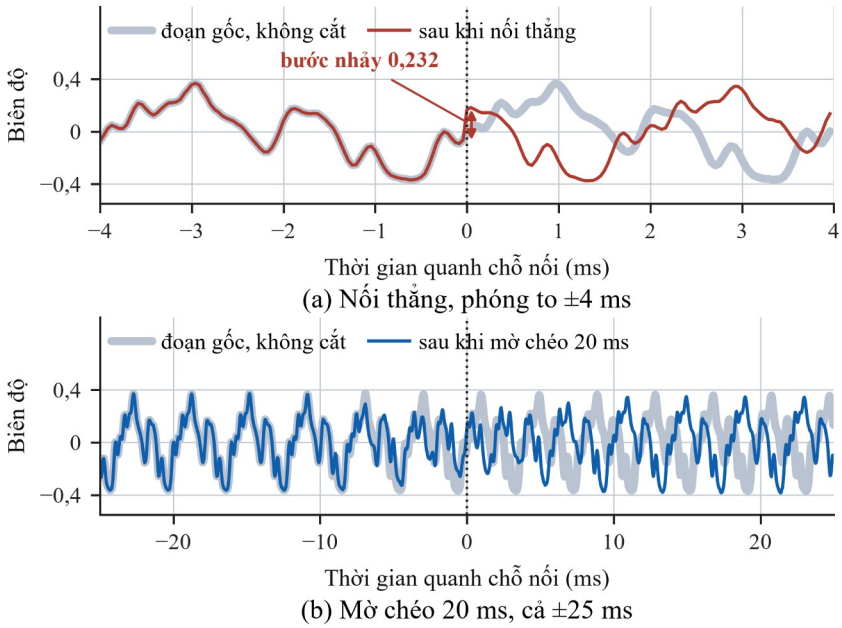
Đo bước nhảy lớn nhất trong một cửa sổ rộng quanh chỗ nối. Cửa sổ rộng thì bắt được cả những tiếng gõ thật của bản nhạc, và với mờ chéo 1, 2 và 5 mili giây thì cả ba cho ra *đúng* cùng một con số — dấu hiệu rõ ràng rằng phép đo đang đo nhầm thứ khác.

Phép đo cuối cùng dựa trên ba nguyên tắc, và cả ba đều đáng nhớ vì chúng áp dụng được cho rất nhiều phép đo khác.

- **Dựng một phép thử có kiểm soát.** Thay vì ghép hai đoạn khác nhau — khi ấy chỗ nối lẫn với chỗ đổi nội dung — hãy cắt *một* đoạn tại vị trí k rồi bỏ đi 1.000 mẫu. Hai nửa vẫn cùng nhạc cụ, cùng mức, cùng giai điệu; chúng chỉ khác nhau đúng một thứ là chỗ nối.
- **Giữ chỗ nối đứng yên.** Nếu mờ chéo dài hơn thì chỗ chuyển dịch về phía trước, và cửa sổ đo trượt theo. Cách sửa là ghép sao cho tâm chỗ chuyển luôn rơi đúng vào mẫu k, dù mờ chéo dài bao nhiêu.
- **Lặp lại ở nhiều vị trí rồi lấy trung vị.** Một chỗ cắt có thể may hoặc rủi: nếu hai mẫu ở hai bên tình cờ gần bằng nhau thì không có bước nhảy nào. Bài chạy phép thử ở 372 vị trí k khác nhau và lấy trung vị.

$$\delta = \max_{n \in W} |y[n+1] - y[n]|, \qquad r = \frac{\delta}{\delta_{\text{nén}}} \tag{14.13}$$

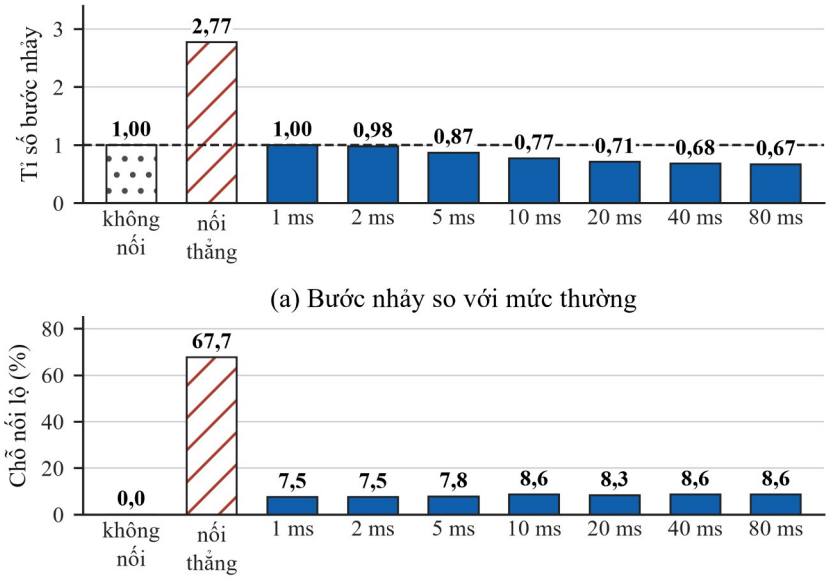
Thước đo là bước nhảy lớn nhất giữa hai mẫu liền nhau trong một cửa sổ hẹp 2 mili giây quanh chỗ nối, chia cho bước nhảy đo ở *đúng* chỗ ấy trên đoạn chưa cắt. Tỉ số bằng 1 nghĩa là chỗ nối không khác gì chỗ bình thường.



(a) Nối thẳng, phóng to ±4 ms

(b) Mờ chéo 20 ms, cả ±25 ms

Hình 14.15 Cùng một chỗ cắt, nối thẳng và nối bằng mờ chéo 20 mili giây



(a) Bước nhảy so với mức thường

(b) Tỉ lệ chỗ nối vượt gấp đôi mức thường

Hình 14.16 Bước nhảy ở mối nối, trung vị trên 372 chỗ cắt

Bảng 14.20 Bước nhảy ở mối nối theo độ dài mờ chéo

Cách nối	Bước nhảy	Gấp mức thường	Tỉ lệ chỗ nối lộ	Cao tần thêm
Không có mối nối	0,100	1,00	0,0 %	0,00 dB
Nối thẳng	0,287	2,77	67,7 %	+10,85 dB
Mờ chéo 1 ms	0,102	1,00	7,5 %	+0,53 dB
Mờ chéo 2 ms	0,096	0,98	7,5 %	+0,60 dB
Mờ chéo 5 ms	0,085	0,87	7,8 %	+0,98 dB
Mờ chéo 10 ms	0,076	0,77	8,6 %	+1,53 dB
Mờ chéo 20 ms	0,071	0,71	8,3 %	+2,04 dB
Mờ chéo 40 ms	0,068	0,68	8,6 %	+2,17 dB
Mờ chéo 80 ms	0,067	0,67	8,6 %	+2,18 dB

Kết quả có hai phần, và phần thứ hai là phần bất ngờ.

Một mili giây đã đủ xoá bậc thang

Nối thẳng cho bước nhảy gấp 2,77 lần mức bình thường, và ở 67,7 phần trăm số chỗ cắt thì bước nhảy ấy vượt gấp đôi mức thường. Đó là tiếng tách nghe rõ.

Chỉ một mili giây mờ chéo đã đưa tỉ số về đúng 1,00 — bằng mức bình thường của bản nhạc. Mờ chéo dài hơn *không* xoá bậc thang tốt hơn; nó chỉ làm tỉ số tụt xuống dưới 1 vì hai nguồn không tương quan trộn lại có độ dốc nhỏ hơn.

Vậy vì sao vẫn dùng 20 đến 40 mili giây? Vì bước nhảy không phải thứ duy nhất tai nghe thấy. Cột cuối cho biết mờ chéo càng dài thì càng nhiều *chồng nội dung*: trong suốt thời gian ấy chúng ta nghe thấy cả hai đoạn cùng lúc. Chọn độ dài là cân giữa hai thứ đó, không phải chỉ giảm bước nhảy.

Quy tắc thực hành rút ra: với nhạc có nhịp rõ, dùng 20 đến 40 mili giây và cắt ở chỗ sát trước một nhịp. Với giọng nói, dùng 5 đến 10 mili giây và cắt ở chỗ lặng giữa hai từ — chồng nội dung trong giọng nói nghe rõ hơn nhiều so với trong nhạc.

14.5.4 Ghép cả ba đoạn và cái giá của việc bỏ sót cắt lặng

Bảng 14.21 Ghép ba đoạn bằng hai luật, mờ chéo 40 mili giây

	Luật tuyến tính	Luật đẳng công suất
Độ dài	22,27 giây	22,27 giây
Độ to tích hợp	−14,03 LUFS	−14,02 LUFS
Dải độ to (LRA)	0,31 *lu*	0,31 *lu*
Đỉnh	−0,56 dBFS	−0,56 dBFS

Trên bản hoàn chỉnh, hai luật gần như không khác nhau ở các con số tổng — vì hai chỗ hõm 3 decibel chỉ kéo dài 40 mili giây trong 22 giây. Đó là lý do phải đo *tại* chỗ nối chứ không đo trên cả bản: thước đo tổng làm loãng đúng thứ cần tìm.

Bảng 14.22 Quên cắt lặng thì bản ghép dài thêm bao nhiêu

	Có cắt lặng	Không cắt lặng
Độ dài bản ghép ba đoạn	22,27 giây	23,65 giây
Phần thừa	—	1,38 giây
Nếu ghép 20 đoạn	—	khoảng 9 giây im lặng rải rác

Một giây ba mươi tám phần trăm giây trên ba đoạn nghe như bản nhạc bị ngắt quãng. Và vì lặng nằm giữa bản chứ không ở hai đầu, không phần mềm nào tự dọn được — phải dọn từng đoạn trước khi ghép.

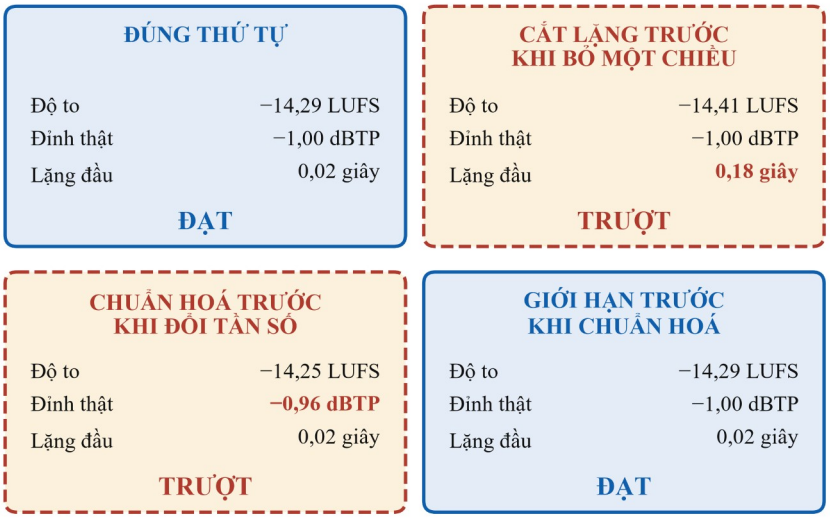
14.6 Đường ống hậu kỳ — thứ tự và ngân sách

14.6.1 Bốn cách xếp cùng năm khối

Năm khối của đường ống — bỏ lệch một chiều, cắt lặng, đổi tần số, chuẩn hoá độ to, giới hạn — xếp được theo rất nhiều thứ tự. Mục này đo bốn thứ tự, trong đó ba thứ tự trông hoàn toàn hợp lý.

Liệt kê 14.8 Đường ống hậu kỳ, đúng thứ tự

```
def duong_ong_hau_ky(x, sr_vao, sr_ra=48000, lufs_dich=-14.0,
                     tran_db=-1.0, so_canh=16, cat_lang=True):
    y = bo_lech_mot_chieu(x, sr_vao)          # 1 — phai dau tien
    if cat_lang:
        y = cat_lang_dau_cuoi(y, sr_vao)      # 2 — sau khi het lech
    if int(sr_ra) != int(sr_vao):
        y = doi_sr(y, sr_vao, sr_ra, so_canh) # 3 — truoc moi phep do
    y, _, _ = chuan_hoa_lufs(y, lufs_dich, tran_db, sr_ra)
    y, _    = gioi_han(y, tran_db, 5.0, 50.0, sr_ra)
    return y
```



Hai cách xếp sai đều cho ra tệp nghe "vẫn được"; chỉ bộ kiểm mới chỉ ra chỗ sai.

Trượt vì: lặng đầu 0,18 giây còn nguyên · đỉnh thật −0,96 dBTP vượt trần −1,0.

Hình 14.17 Bốn cách xếp cùng năm khối, và kết quả

Bảng 14.23 Bốn thứ tự, đo trên đoạn 1

Thứ tự	Độ to	Đỉnh thật	Lặng đầu	Kết luận
Đúng thứ tự	−14,29 LUFS	−1,00 dBTP	0,02 giây	*Đạt*
Cắt lặng trước khi bỏ một chiều	−14,41 LUFS	−1,00 dBTP	0,18 giây	*Trượt*
Chuẩn hoá trước khi đổi tần số	−14,25 LUFS	−0,96 dBTP	0,02 giây	*Trượt*
Giới hạn trước khi chuẩn hoá	−14,29 LUFS	−1,00 dBTP	0,02 giây	*Đạt*

Hai cách sai, hai lý do khác nhau

Cắt lặng trước khi bỏ một chiều: máy dò lặng đọc 0 mili giây vì lệch làm mọi khối đều kêu, nên nó không cắt gì cả. Bản giao còn nguyên 0,18 giây lặng ở đầu. Đây đúng là điểm dễ sai ở mục 14.2.4.

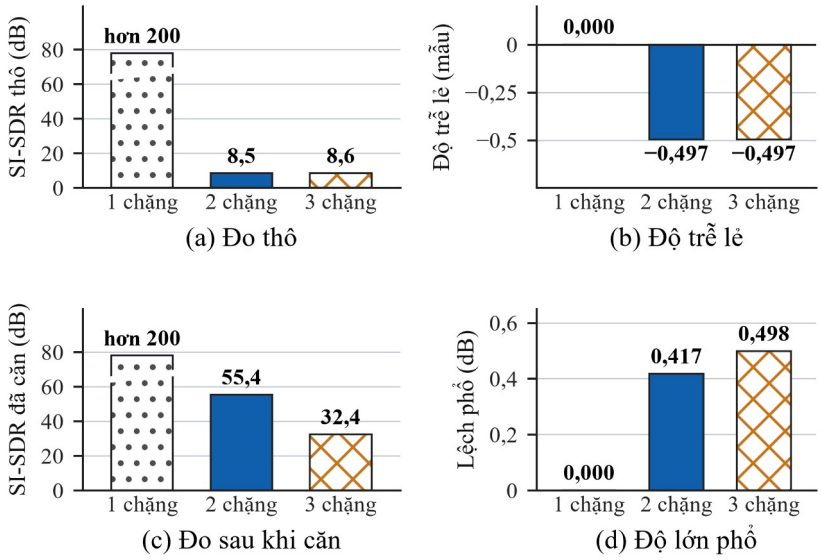
Chuẩn hoá trước khi đổi tần số: đỉnh thật đo được ở 24 kHz không phải đỉnh thật của bản 48 kHz. Bộ giới hạn đã làm đúng việc của nó ở 24 kHz, nhưng phép đổi tần số sau đó đẩy đỉnh thật lên −0,96 dBTP, vượt trần −1,0.

Cách thứ tư — giới hạn trước khi chuẩn hoá — vẫn đạt, nhưng nó làm bộ giới hạn phải hạ gain ở nhiều chỗ hơn cần thiết. Nó không sai, chỉ là làm việc thừa.

14.6.2 Đổi tần số nhiều chặng: một phép đo suýt nữa cho kết quả sai

Một câu hỏi thực tế: nếu tệp đã đi qua vài lần đổi tần số — 24 kHz lên 44,1 kHz rồi lên 48 kHz — thì nó sai lệch thêm bao nhiêu so với đổi thẳng một chặng?

Đo bằng SI-SDR, thước đo đã dùng suốt chương 13, cho một con số gây bất ngờ.



Hình 14.18 Ba đường đổi tần số, đo thô và đo sau khi căn

Bảng 14.24 Đổi tần số nhiều chặng, ba cách đo

Đường đi	SI-SDR đo thô	Độ trễ lẻ	SI-SDR sau khi căn	Lệch độ lớn phổ
1 chặng: 24 → 48	hơn 200 dB	0,000 mẫu	hơn 200 dB	0,000 dB
2 chặng: 24 → 44,1 → 48	8,5 dB	−0,497 mẫu	55,4 dB	0,417 dB
3 chặng: 24 → 44,1 → 22,05 → 48	8,6 dB	−0,497 mẫu	32,4 dB	0,498 dB

Tám phẩy năm decibel là một con số sai

SI-SDR so từng mẫu một. Nếu hai tín hiệu hoàn toàn giống nhau nhưng lệch nhau nửa mẫu thì nó phạt rất nặng — và đúng nửa mẫu là thứ mà chuỗi hai lần đổi tần số để lại. Bằng chứng nằm ở hai cột sau. Độ trễ lẻ đo được là −0,497 mẫu, gần đúng một nửa. Dịch tín hiệu lại đúng nửa mẫu ấy rồi đo lại thì SI-SDR nhảy từ 8,5 lên 55,4 decibel.

Và thước đo không quan tâm pha — chênh lệch *độ lớn* phổ — cho trung vị 0,0001 decibel trên cả dải. Nói cách khác: đi vòng qua 44,1 kHz gần như không làm sai lệch gì cả.

Ai tin con số 8,5 decibel sẽ đi chữa một lỗi không tồn tại. Đây là lần thứ tư trong cuốn sách một thước đo khách quan nói ra điều không đúng với tai — ba lần trước nằm ở mục 13.8.

$$|E(f)| \approx 2\pi f \Delta t\, |X(f)| \tag{14.14}$$

Công thức giải thích vì sao sai số do trễ lẻ tăng theo tần số: dịch một khoảng delta t làm pha xoay tỉ lệ với tần số, nên biên độ sai khác cũng tỉ lệ với tần số. Đó chính là hình dạng quan sát được khi vẽ sai số theo thời gian của một tín hiệu quét tần.

Nhưng đường ba chặng thì sai thật: 32,4 decibel sau khi căn, vì nó đi qua 22,05 kHz và mất hẳn phần dải trên 11 kHz. Đó là mất mát thật, không phải ảo giác của thước đo.

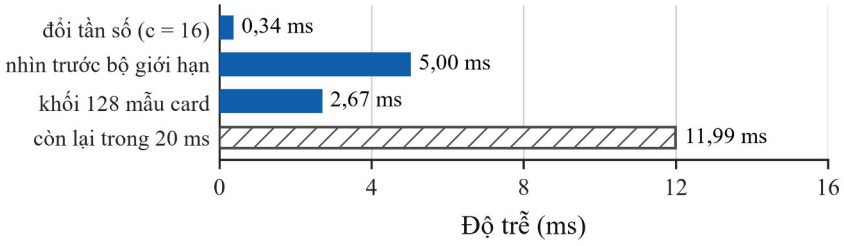
Liệt kê 14.9 Đo độ trễ lẻ và căn lại trước khi so sánh

```
tre = tre_le_giua_hai(mot_chang, nhieu_chang, sr, 200.0, 8000.0)
z   = dich_le_mau(nhieu_chang, -tre)      # can lai bang xoay pha
print(si_sdr_db(z, mot_chang))            # 55.4 thay vi 8.5
print(lech_do_lon_pho_db(mot_chang, nhieu_chang, sr, 50, 9000))
# (0.417, 0.0001) — lech lon nhat va lech trung vi, don vi dB
```

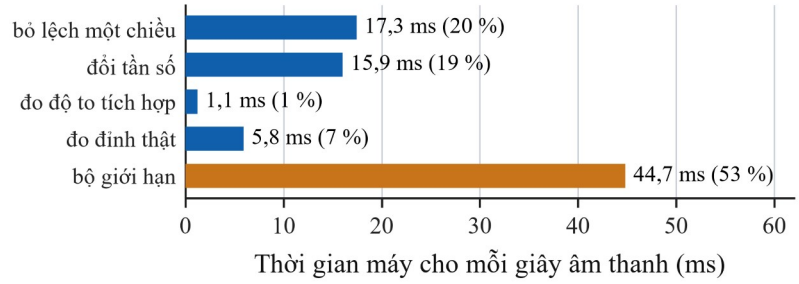
14.6.3 Hai ngân sách: độ trễ và thời gian máy

Đường ống có hai loại chi phí, và nhầm lẫn giữa chúng là một lỗi thường gặp khi thiết kế. *Độ trễ* quyết định đường ống có dùng được khi chơi trực tiếp hay không. *Thời gian máy* quyết định dựng một bản dài mất bao lâu. Chúng hoàn toàn độc lập với nhau.

$$\tau_{\text{tổng}} = \sum_i \tau_i , \qquad \tau_{\text{còn}} = \tau_{\text{hạn}} - \tau_{\text{tổng}} \tag{14.15}$$



(a) Độ trễ: ba khối dùng 8,01 ms, còn dư 11,99 ms



(b) Thời gian máy: tổng 84,8 ms, nhanh gấp 11,8 lần thời gian thực

Hình 14.19 Hai ngân sách của cùng một đường ống

Bảng 14.25 Ngân sách độ trễ cho một đường ống chơi trực tiếp

Khối	Độ trễ	Đổi được không
Đổi tần số 24 → 48 kHz, c = 16	0,34 ms	Có — nhưng đã rất nhỏ vì L = 2
Nhìn trước của bộ giới hạn	5,00 ms	Có — đổi lấy khả năng bắt gai
Khối 128 mẫu của card âm thanh	2,67 ms	Có — đổi lấy tải máy
Tổng	8,01 ms	—
Còn lại trong ngân sách 20 ms	11,99 ms	—

Nếu đường đổi là 44,1 lên 48 kHz thay vì 24 lên 48 kHz thì riêng bộ lọc đã tốn 50,5 mili giây với c bằng 16 — vượt ngân sách 20 mili giây gấp hai lần rưỡi. Đó là một lý do nữa để chọn tần số mô hình có tỉ số đơn giản.

Bảng 14.26 Thời gian máy, đo trên mỗi giây âm thanh

Khối	Thời gian	Phần trăm của tổng
Bỏ lệch một chiều	17,3 ms	20 %
Đổi tần số	15,9 ms	19 %
Đo độ to tích hợp	1,1 ms	1 %
Đo đỉnh thật	5,8 ms	7 %
Bộ giới hạn	44,7 ms	53 %
Tổng	84,8 ms	100 %

Cả đường ống chạy nhanh gấp 11,8 lần thời gian thực — dựng một bản ba phút mất khoảng mười lăm giây. Với hậu kỳ thì quá đủ; với một plugin chạy trực tiếp thì bản Python thuần này không dùng được, và phần tốn nhất là bộ giới hạn.

Đọc bảng thời gian máy để làm gì

Nó chỉ ra chỗ đáng tối ưu. Bộ giới hạn chiếm hơn một nửa, nên nếu cần nhanh gấp đôi thì viết lại riêng nó là đủ — tối ưu bốn khối còn lại cộng lại cũng không bằng.

Nó cũng chỉ ra chỗ *không* đáng tối ưu. Đo độ to chiếm 1 phần trăm; viết lại nó bằng C cũng chỉ tiết kiệm 1 mili giây mỗi giây.

Đây là thói quen nên mang sang mọi việc khác: đo trước, rồi mới tối ưu. Hầu như lần nào trực giác cũng chỉ sai chỗ.

14.7 Đo và kiểm tự động

14.7.1 Bộ kiểm bàn giao: một lệnh, bảy dòng

Nghe lại toàn bộ bản nhạc trước mỗi lần giao là việc tốn thời gian và không đáng tin — tai mệt dần, và người làm ra bản nhạc là người nghe nó kém khách quan nhất. Thứ thay cho việc ấy là một hàm trả về bảng đạt hay không đạt.

Liệt kê 14.10 Bộ kiểm bàn giao

```
def kiem_ban_giao(x, sr=SR, lufs_dich=-14.0, sai_so_lufs=0.5,
                  tran_tp=-1.0, lech_toi_da=0.001,
                  lang_toi_da_ms=250.0, cat_toi_da=0.0,
                  sr_goc_it_nhat=None):
    lufs = do_to_tich_hop(x, sr)
    tp   = true_peak_dbtp(x)
    lech = abs(lech_mot_chieu(x))
    dau_ms, cuoi_ms = [1000.0*v for v in lang_dau_cuoi(x, sr)]
    cat  = ti_le_mau_cat(x)
    # ... moi tieu chi tra ve (ten, so do, nguong, dat)
```

Tiêu chí	Ngưỡng	Bản thô	Đạt?	Sau ống	Đạt?
độ to tích hợp (LUFS)	−14,0 ± 0,5	**−17,76**	**trượt**	**−14,29**	đạt
đỉnh thật (dBTP)	≤ −1,0	−4,47	đạt	**−1,00**	đạt
lệch một chiều	≤ 0,0010	**0,0122**	**trượt**	**0,0000**	đạt
lặng đầu (ms)	≤ 250	0	đạt	**20**	đạt
lặng cuối (ms)	≤ 250	0	đạt	**0**	đạt
mẫu chạm trần (%)	≤ 0,000	0,000	đạt	**0,000**	đạt
vách phổ (Hz)	≥ 12.000	12.000	đạt	**13.078**	đạt

Đoạn 1: bản thô đạt 5/7, sau đường ống đạt 7/7.

Dòng lặng đầu của bản thô đọc 0 ms và được tính là đạt, nhưng kết quả này sai: lệch một chiều làm mọi khối đều có tín hiệu nên máy dò không thấy lặng. Một tiêu chí trượt có thể che một tiêu chí khác.

Hình 14.20 Bộ kiểm bàn giao chạy trên đoạn 1

Bảng 14.27 Bảy tiêu chí và lý do của từng ngưỡng

Tiêu chí	Ngưỡng	Ngưỡng ấy từ đâu ra
Độ to tích hợp	−14,0 ± 0,5 LUFS	Mức đích của nơi nhận; sai số 0,5 *lu* nằm dưới ngưỡng nghe ra
Đỉnh thật	không quá −1,0 dBTP	Mục 14.4.4: trần 0 dBFS còn để vượt +0,02 dBTP sau khi đổi tần số
Lệch một chiều	không quá 0,001	Dưới mức này thì máy dò lặng và bộ nén không bị ảnh hưởng
Lặng đầu	không quá 250 ms	Đủ cho một khoảng lấy hơi, không đủ để nghe như đứt quãng
Lặng cuối	không quá 250 ms	Như trên
Mẫu chạm trần	0,000 %	Một mẫu chạm trần là một chỗ méo; không có lý do để dung thứ
Vách phổ	không thấp hơn Nyquist của tần số gốc	Phát hiện tệp đã bị đi xuống rồi nâng lên, hoặc qua codec

Bảng 14.28 Bộ kiểm chạy trên ba đoạn, trước và sau đường ống

Đoạn	Bản thô	Sau đường ống
1	5 / 7	7 / 7
2	5 / 7	6 / 7
3	5 / 7	6 / 7

Đoạn 2 và đoạn 3 còn trượt một tiêu chí sau đường ống: vách phổ. Lý do là chúng nhỏ hơn nên phần cao của phổ chìm sâu hơn dưới nền nhiễu, và máy dò vách đọc ra một con số hơi thấp. Đó không phải lỗi của đường ống mà là giới hạn của phép dò — và một bộ kiểm trung thực phải để lộ cả những chỗ chính nó không chắc.

14.7.2 Một tiêu chí trượt có thể che một tiêu chí khác

Bảng kiểm trên đoạn thô có một chi tiết đáng dừng lại. Dòng "lặng đầu" đọc 0 mili giây và *đạt* — trong khi đoạn ấy có 180 mili giây lặng thật.

Đọc bảng kiểm theo thứ tự, không đọc rời từng dòng

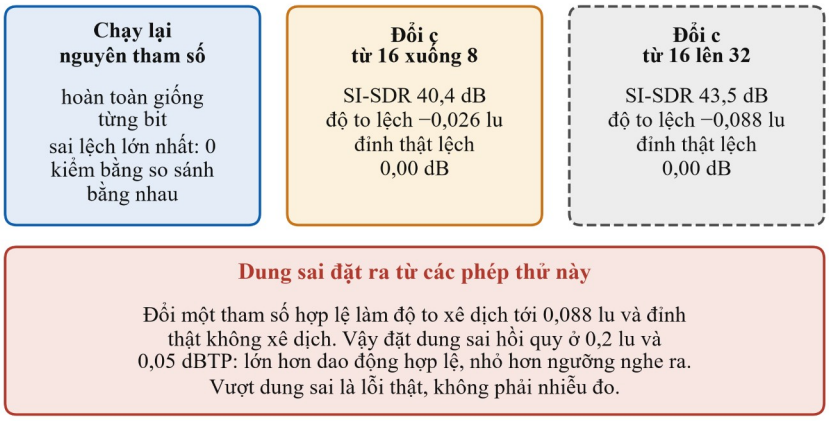
Dòng "lệch một chiều" trượt. Mà lệch một chiều là thứ làm máy dò lặng mù. Nên dòng "lặng đầu" đạt *không* có nghĩa là không có lặng — nó chỉ có nghĩa là phép đo ấy hiện không đáng tin.

Quy tắc đọc: khi một tiêu chí nền tảng trượt, mọi tiêu chí phụ thuộc vào nó đều phải coi là chưa đo được. Sửa tiêu chí nền tảng rồi chạy lại toàn bộ.

Trong bộ kiểm này, thứ tự phụ thuộc là: lệch một chiều đứng trước lặng hai đầu; tần số lấy mẫu đứng trước đỉnh thật; và cả hai đứng trước độ to.

14.7.3 Kiểm hồi quy: biết trước cái gì được phép đổi

Bộ kiểm bàn giao trả lời câu "tệp này có đạt không". Còn một câu hỏi khác, và nó quan trọng không kém khi một nhóm cùng sửa mã: "lần sửa vừa rồi có làm đổi kết quả không, và đổi bao nhiêu thì là quá nhiều".



Hình 14.21 Ba phép thử để đặt ra dung sai hồi quy

Bảng 14.29 Ba phép thử hồi quy trên cùng một đoạn

Phép thử	SI-SDR	Lệch độ to	Lệch đỉnh thật
Chạy lại y nguyên	hoàn toàn giống từng bit	0,000 *lu*	0,00 dB
Đổi c từ 16 xuống 8	40,4 dB	−0,026 *lu*	0,00 dB
Đổi c từ 16 lên 32	43,5 dB	−0,088 *lu*	0,00 dB

Dòng đầu là dòng quan trọng nhất: chạy lại cùng một đầu vào với cùng tham số cho ra kết quả *hoàn toàn giống từng bit*. Đường ống không có chỗ nào ngẫu nhiên, nên phép kiểm hồi quy đơn giản nhất là so sánh bằng nhau.

Hai dòng sau đặt ra dung sai. Đổi một tham số hợp lệ làm độ to xê dịch tới 0,088 *lu* và không làm đỉnh thật xê dịch chút nào. Vậy đặt dung sai ở 0,2 *lu* và 0,05 dBTP: lớn hơn dao động hợp lệ, nhỏ hơn ngưỡng nghe ra.

> **Vì sao phải đo dung sai chứ không đoán nó**
>
> Dung sai đặt quá chặt thì mỗi lần sửa mã hợp lệ đều báo lỗi, và sau ba lần như thế cả nhóm sẽ bỏ qua bộ kiểm.
>
> Dung sai đặt quá rộng thì một lỗi thật đi qua. Con số 0,2 *lu* ở đây không bốc ra từ cảm tính: nó gấp hơn hai lần dao động đo được, và bằng một phần ba ngưỡng nghe ra.
>
> Cách làm này — đo dao động của những thay đổi *hợp lệ* rồi đặt ngưỡng ở giữa dao động ấy và ngưỡng có ý nghĩa — dùng được cho gần như mọi bộ kiểm số học, không riêng âm thanh.

14.8 Dự án thực hành — giai đoạn 4

14.8.1 Vị trí của giai đoạn 4 trong dự án thực hành

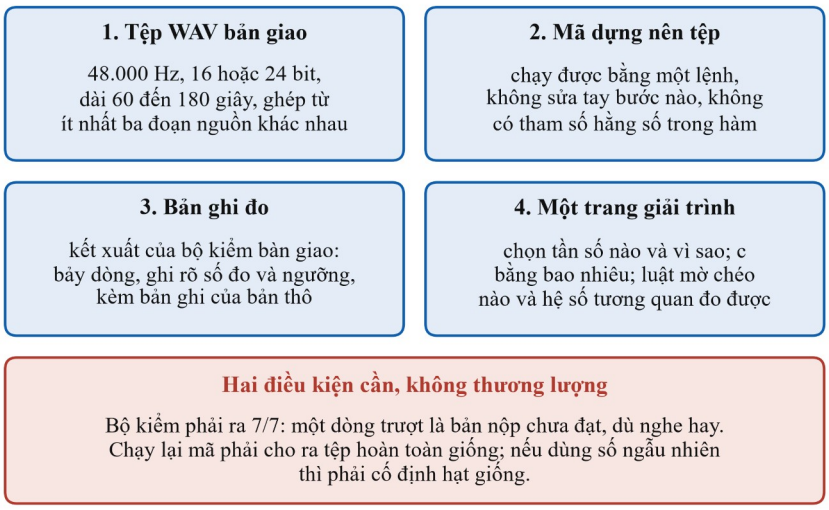
Dự án thực hành xuyên suốt cuốn sách đi qua năm giai đoạn. Ba giai đoạn đầu ứng với các chương trước: dựng được tín hiệu và đo được nó, dựng được chuỗi hiệu ứng, và dựng được bản trộn nhiều lớp chạy trên trình duyệt. Giai đoạn 4 là giai đoạn biến thứ đang có thành một tệp *giao được*.

Bảng 14.30 Năm giai đoạn của dự án thực hành

Giai đoạn	Mốc	Chương	Trạng thái
1	Sinh và đo tín hiệu	Chương 4	Làm sau chương 4
2	Chuỗi hiệu ứng có đo	Chương 9	Làm sau chương 9
3	Bản trộn chạy trên trình duyệt	Chương 12	Làm sau chương 12
4	Bản giao đạt bộ kiểm	Chương 14	Làm sau chương này
5	Lớp bảo vệ, tài liệu và trình bày hoàn chỉnh	Chương 15	Làm sau chương 15

Giai đoạn 4 khác ba giai đoạn trước ở một chỗ: nó không yêu cầu thêm một kỹ thuật mới nào. Nó yêu cầu ghép những thứ đã có thành một quy trình chạy được bằng một lệnh, và chứng minh bằng số rằng kết quả đạt chuẩn.

14.8.2 Kết quả cần có



Hình 14.22 Bốn kết quả cần có ở giai đoạn 4

Bảng 14.31 Bốn kết quả cần có và tiêu chuẩn của từng thứ

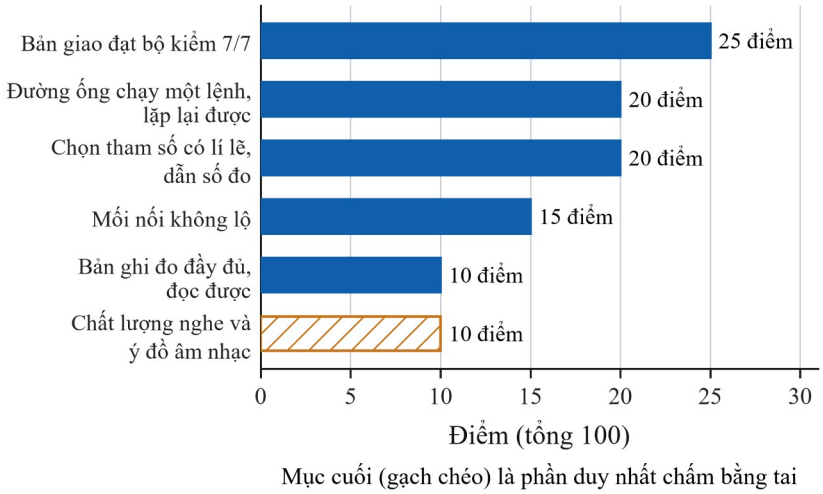
Kết quả	Tiêu chuẩn
Tệp WAV bản giao	48.000 Hz, 16 hoặc 24 bit, dài 60 đến 180 giây, ghép từ ít nhất ba đoạn nguồn khác nhau
Mã dựng nên nó	Chạy được bằng một lệnh từ thư mục nguồn tới tệp giao, không sửa tay bước nào; không có tham số nào hằng số lẫn trong thân hàm
Bản ghi đo	Kết xuất của bộ kiểm bàn giao, bảy dòng, ghi rõ số đo và ngưỡng; kèm bản ghi đo của bản *thô* để thấy đường ống đã sửa được gì
Một trang giải trình	Chọn tần số nào và vì sao; c bằng bao nhiêu và dựa vào số đo nào; luật mờ chéo nào và hệ số tương quan đo được là bao nhiêu

Hai điều kiện cần — không thương lượng

Bộ kiểm phải ra 7 trên 7. Một dòng trượt là bản giao chưa đạt, dù nghe hay đến đâu. Nếu có lý do chính đáng để một tiêu chí không áp dụng được — ví dụ nguồn thật sự là 16 kHz — thì phải nêu trong trang giải trình và đổi ngưỡng của tiêu chí ấy một cách có căn cứ, chứ không được bỏ dòng ấy đi.

Chạy lại mã phải cho ra tệp *hoàn toàn giống*. Nếu mỗi lần chạy một khác thì không có cách nào kiểm được, và cũng không có cách nào sửa lỗi. Nếu có dùng số ngẫu nhiên ở đâu thì phải cố định hạt giống.

14.8.3 Tự đánh giá kết quả



Hình 14.23 Các hạng mục tự đánh giá giai đoạn 4

Khi tự đánh giá giai đoạn này, hãy kiểm bằng số đo trước: bản giao đạt bộ kiểm 7 trên 7, đường ống chạy một lệnh và lặp lại được, mỗi lựa chọn tham số dẫn ít nhất một con số, mối nối không lộ, và bản ghi đo đầy đủ cả trước lẫn sau. Phần nghe để sau cùng. Đó không phải để hạ thấp phần nghe, mà vì phần nghe chỉ có ý

nghĩa khi phần kỹ thuật đã sạch — một bản nhạc hay nhưng bị cắt đỉnh ở ba chỗ và có hai chỗ nối lộ thì không ai nghe được phần hay của nó.

14.8.4 Ba lỗi khiến bản giao trượt bộ kiểm nhiều nhất

Ba lỗi dự đoán trước được

Quên rằng mọi phép đo phải ở *cuối* đường ống. Rất nhiều người đo đỉnh thật ngay sau bộ giới hạn rồi đổi tần số, và ngạc nhiên khi bộ kiểm báo vượt trần. Mục 14.6.1 đã đo đúng trường hợp này.

Ghép các đoạn mà không đo hệ số tương quan, rồi dùng luật tuyến tính vì nó "tự nhiên hơn". Kết quả là một chỗ hõm ba decibel tại mỗi mối nối. Đo trước, chọn sau.

Để sót một tham số hằng số trong thân hàm — thường là tần số lấy mẫu 48000 viết thẳng vào giữa một phép tính. Khi ấy đổi tần số đích sẽ cho kết quả sai một cách âm thầm. Mọi tham số phải nằm ở một chỗ duy nhất.

14.9 Thực hành, bài tập và tổng kết

14.9.1 Ba bài thực hành

Ba bài dưới đây làm ngay sau khi đọc xong chương, mỗi bài khoảng mười lăm phút, và cả ba đều chạy trên vật liệu do mô-đun của cuốn sách sinh ra nên không cần tệp ngoài.

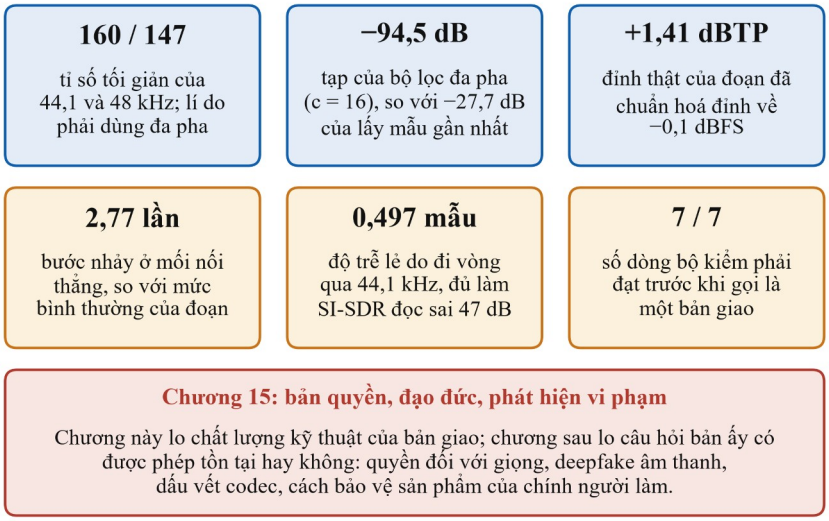
- **Thực hành 1 — sáu chỉ số đầu tiên.** Sinh ba đoạn bằng hàm của chương, đo sáu chỉ số ở mục 14.2.2 trên từng đoạn, rồi lặp lại phép đo *sau* khi bỏ lệch một chiều. Chỉ ra chỉ số nào đổi và giải thích vì sao.
- **Thực hành 2 — ba cách đổi tần số.** Đưa một hình sin 997 Hz từ 44,1 lên 48 kHz bằng ba cách, vẽ phổ của cả ba, và đo phần năng lượng nằm ngoài 997 Hz. Đối chiếu với bảng ở mục 14.3.4.
- **Thực hành 3 — hai luật mờ chéo.** Ghép hai đoạn nhiễu trắng độc lập bằng cả hai luật, đo mức ở giữa chỗ chuyển, rồi lặp lại với hai đoạn hoàn toàn giống nhau. Bốn con số phải khớp bảng bốn ô ở mục 14.5.2 trong phạm vi 0,2 decibel.

14.9.2 Bảy bài tập

- **Bài 1 — dò vách phổ.** Lấy ba tệp nhạc bất kỳ trên máy của người đọc, dò vách phổ của từng tệp, và đoán xem chúng đã qua những khâu nào. Nếu tìm được một tệp có vách ở khoảng 15,5 kHz thì giữ lại — chương 15 sẽ dùng tới.
- **Bài 2 — đường cong số cạnh.** Dựng lại bảng ở mục 14.3.5 cho đường 48 xuống 24 kHz thay vì 44,1 lên 48 kHz. Chỉ ra chỗ khác biệt lớn nhất so với bảng trong bài, và giải thích.
- **Bài 3 — chứng minh bằng số rằng đi xuống không đảo ngược được.** Lấy một đoạn 48 kHz, đưa xuống 24 kHz rồi nâng lại 48 kHz, và đo SI-SDR so với bản gốc theo hai cách: trên tín hiệu đầy dải, và trên tín hiệu chỉ có nội dung dưới 9,6 kHz. Hai con số ấy khác nhau rất nhiều — nói rõ vì sao.

- **Bài 4 — đỉnh thật của chính người đọc.** Lấy năm tệp nhạc trên máy, đo đỉnh mẫu và đỉnh thật của từng tệp, và đếm xem bao nhiêu tệp có đỉnh thật vượt 0 dBTP. Kết quả thường làm người đo ngạc nhiên.
- **Bài 5 — chọn độ dài mờ chéo cho giọng nói.** Lặp lại phép đo ở mục 14.5.3 nhưng trên một đoạn giọng nói thay vì nhạc. Bảng kết quả sẽ khác — chỉ ra chỗ khác và đưa ra khuyến nghị độ dài riêng cho giọng nói.
- **Bài 6 — thêm một tiêu chí vào bộ kiểm.** Thêm một tiêu chí thứ tám vào bộ kiểm bàn giao. Tiêu chí ấy phải có một con số đo được, một ngưỡng, và một lập luận cho ngưỡng ấy theo đúng cách ở mục 14.7.3. Gợi ý: dải độ to, hoặc cân bằng trái phải, hoặc tương quan giữa hai kênh.
- **Bài 7 — điểm dễ sai của thước đo.** Dựng lại phép đo ở mục 14.6.2 trên vật liệu của chính người đọc. Đo SI-SDR thô, đo độ trễ lẻ, căn lại rồi đo lần nữa. Viết một đoạn ngắn nói xem nếu chỉ nhìn con số đầu tiên thì sẽ kết luận sai như thế nào.

14.9.3 Tổng kết chương 14



Hình 14.24 Chương 14 gói lại trong sáu con số

Chương này không có một khái niệm nào thật sự mới. Toàn bộ giá trị của nó nằm ở chỗ khác: nó cho thấy rằng cùng những khối ấy, xếp sai thứ tự thì cho ra một tệp không đạt, và rằng chỉ có phép đo mới phân biệt được hai trường hợp — tai thì không.

Bốn điều đáng mang ra khỏi chương này

Mọi phép đo và mọi phép đặt mức phải nằm ở *cuối* đường ống. Con số chúng ta hứa với người nhận là con số của tệp giao, không phải của một bản trung gian.

Một tiêu chí trượt có thể *che* một tiêu chí khác. Đọc bảng kiểm theo thứ tự phụ thuộc, đừng đọc rời từng dòng.

Thước đo có thể cho kết luận sai một cách rất thuyết phục. Nửa mẫu trễ làm SI-SDR đọc sai 47 decibel, trong khi phổ gần như không đổi. Trước khi tin một con số kém, hãy hỏi xem thước đo ấy đang phạt cái gì.

Dung sai phải *đo* ra, không đoán ra: đo dao động của những thay đổi hợp lệ, rồi đặt ngưỡng giữa dao động ấy và ngưỡng có ý nghĩa.

Chương 15 — chương cuối cùng của cuốn sách — chuyển sang một loại câu hỏi khác hẳn. Chương này lo chất lượng kỹ thuật của bản giao; chương sau lo việc bản ấy có được phép tồn tại không: quyền đối với giọng, deepfake âm thanh, dấu vết codec dùng để tìm vi phạm, và cách bảo vệ chính sản phẩm của mình.

Tài liệu tham khảo của chương

Nguồn trực tiếp của nội dung trong bài

- ITU-R BS.1770-4 — Algorithms to measure audio programme loudness and true-peak audio level. Nguồn của phép đo độ to tích hợp và của định nghĩa đỉnh thật ở công thức (14.10), kể cả yêu cầu lấy mẫu quá ít nhất bốn lần.
- EBU Tech 3341 và Tech 3342 — hai tài liệu đi kèm R 128, định nghĩa đồng hồ đo độ to và dải độ to LRA. Đọc để biết vì sao ngoài một con số LUFS còn cần một con số LRA.
- Crochiere, R. E., Rabiner, L. R. — Multirate Digital Signal Processing. Sách gốc của phép đổi tần số bằng bộ lọc đa pha; toàn bộ mục 14.3.3 là một trường hợp riêng của chương 3 sách này.
- Smith, J. O. — Digital Audio Resampling Home Page, và bộ mã libresample đi kèm. Đây là nguồn của cách dựng bộ lọc bằng cửa sổ Kaiser nhân sinc, và của chính tham số "số cạnh" dùng trong bài.
- Kaiser, J. F. — Nonrecursive Digital Filter Design Using the I0-sinh Window Function. Nguồn của cửa sổ Kaiser và của quan hệ giữa tham số beta với độ triệt dải chặn.
- Le Roux, J. và cộng sự — SDR: Half-Baked or Well Done? Đã dẫn ở chương 13; ở chương này nó là lời cảnh báo cho đúng điểm dễ sai độ trễ lẻ ở mục 14.6.2.

Sách và tài liệu nền

- Oppenheim, A. V., Schafer, R. W. — Discrete-Time Signal Processing, chương về hệ thống nhiều tốc độ. Nền lý thuyết đầy đủ cho công thức (14.3) đến (14.6).

- Katz, B. — Mastering Audio: The Art and the Science. Sách nền cho phần mức giao hàng và trần giới hạn; chương về đỉnh thật đáng đọc trước khi đặt bất kỳ trần nào.
- Zölzer, U. (chủ biên) — DAFX: Digital Audio Effects. Phần về mờ chéo và về ghép đoạn bổ sung cho mục 14.5.
- Smith, S. W. — The Scientist and Engineer's Guide to Digital Signal Processing. Chương về bộ lọc cửa sổ là cách vào nhẹ nhàng nhất cho mục 14.3.3.

Công cụ nên biết

- libsamplerate (Secret Rabbit Code) và soxr. Hai thư viện đổi tần số chất lượng cao; đọc phần mô tả các mức chất lượng của chúng và đối chiếu với bảng số cạnh ở mục 14.3.5.
- ffmpeg với bộ lọc aresample và loudnorm. Công cụ dòng lệnh làm được gần hết đường ống của chương này; dùng nó để đối chiếu với bản tự viết, đừng dùng nó thay cho việc tự viết.
- pyloudnorm và essentia. Hai thư viện Python cài BS.1770; chạy song song với hàm của cuốn sách để kiểm chéo.
- Bộ tự kiểm của mô-đun đi kèm sách. Ba mươi hai hàm mới của chương này đều có phép kiểm riêng, và người đọc chạy chính bộ ấy trên mã của mình để tự đối chiếu.

Đọc thêm theo hướng

- **Muốn đi sâu về đổi tần số:** tìm đọc về bộ lọc nửa băng và về chuỗi nhiều tầng. Khi tỉ số là luỹ thừa của 2, chia phép đổi thành nhiều tầng nửa băng rẻ hơn hẳn một tầng duy nhất.
- **Muốn đi sâu về độ to:** tìm đọc về mức đích của từng nền tảng phát nhạc và về cách chúng chuẩn hoá khi phát. Đó là lý do thật sự khiến cuộc chạy đua độ to kết thúc.
- **Muốn đi sâu về kiểm tự động:** tìm đọc về kiểm hồi quy dựa trên ảnh chụp và về kiểm dựa trên tính chất. Bộ kiểm ở mục 14.7 là dạng đơn giản nhất của loại thứ nhất.
- **Muốn đi sâu về mối nối:** tìm đọc về thuật toán WSOLA và về phép ghép đồng bộ theo chu kỳ cơ bản. Chúng chọn chỗ cắt sao cho hai bên khớp pha nhau, thay vì che chỗ lệch bằng mờ chéo.

Một lưu ý về các con số trong bài

Cách đọc các con số

Các con số về tỉ số hữu tỉ, số phép nhân, độ trễ bộ lọc và dải giữ được là kết quả tính chính xác từ công thức. Chúng đúng với mọi máy và mọi vật liệu.

Các con số về độ triệt ảnh, gợn dải thông, LUFS, đỉnh thật, bước nhảy ở mối nối và SI-SDR là kết quả *đo* trên vật liệu sinh bằng chính mô-đun của cuốn sách, nên chạy lại được hoàn toàn như cũ. Đổi vật liệu thì con số đổi theo; *hình dạng* của kết quả thì không.

Riêng thời gian chạy ở mục 14.6.3 phụ thuộc vào máy. Con số trong bài đo trên một máy ảo phổ thông chạy Python thuần; tỉ lệ giữa các khối thì ổn định hơn nhiều so với con số tuyệt đối.

Phép đo bước nhảy ở mối nối lấy *trung vị* trên 372 vị trí cắt khác nhau, chứ không lấy một vị trí. Lý do và hai phép đo sai đã thử trước đó đều nêu ngay tại mục 14.5.3.

Vật liệu dùng trong bài là tiếng *tổng hợp* mang năm khuyết tật cố ý, không phải bản thu thật. Điều đó làm cho mọi phép đo lặp lại được. Bài tập 4 và bài tập 7 yêu cầu chạy lại trên vật liệu của chính người đọc.

Mã Python in trong bài là mã đã chạy, không phải mã giả.