MỤC LỤC

Chương 9. Hiệu ứng phần 2.....5
9.1 Vị trí của chương trong cuốn sách.....6
9.2 Trễ phân số và họ điều biến.....8
9.3 Bộ lọc bậc hai — biquad.....12
9.4 Cân bằng tần số.....17
9.5 Wah — bộ lọc có tần số chạy.....20
9.6 Điều chế biên độ.....22
9.7 Bộ vang.....25
9.8 Tích chập đáp ứng xung.....29
9.9 Thực hành, bài tập và tổng kết.....31

DANH MỤC HÌNH VẼ

Hình 9.1 Sáu mục kỹ thuật của chương 9.....6
Hình 9.2 Chương 8 chỉ tạo được răng lược; chương 9 chọn được thẳng.....6
Hình 9.3 Mọi giá trị cần lấy đều nằm giữa hai mẫu đã có.....8
Hình 9.4 Nội suy tuyến tính mất cao tần, và mức mất ấy THAY ĐỔI.....9
Hình 9.5 Hồi tiếp làm rãnh sâu hơn, không làm rãnh nhiều hơn.....10
Hình 9.6 Flanger, chorus và vibrato — khác nhau chỉ ở đường độ trễ.....11
Hình 9.7 Bảng chọn tham số cho ba hiệu ứng điều biến.....11
Hình 9.8 Dạng trực tiếp II chuyển vị, và đáp ứng xung của bộ lọc.....13
Hình 9.9 Bảy kiểu biquad và việc của từng kiểu.....14
Hình 9.10 Bảy kiểu chia thành bốn việc, nhìn trên đáp ứng biên độ.....15
Hình 9.11 Ba cách nhìn cùng một con số Q.....15
Hình 9.12 Kệ thấp và kệ cao, bốn lượng nâng mỗi loại.....17
Hình 9.13 Lọc đỉnh, và cách đổi Q sang quãng tám.....17
Hình 9.14 Bốn dải nối tiếp, và phép kiểm trên tín hiệu thật.....18
Hình 9.15 Hai cách làm tiếng sáng hơn, và cái giá ở cột đỉnh.....19
Hình 9.16 Wah là một bộ thông dải có tần số trung tâm chạy.....20
Hình 9.17 Ba giá trị Q trên cùng một đoạn guitar.....21
Hình 9.18 Auto-wah: tần số bộ lọc đi theo đường bao mức.....22
Hình 9.19 Ba dạng dao động, và cách độ sâu đổi sang decibel.....23
Hình 9.20 Sóng mang biến mất; chỉ còn hai dải bên ở hiệu và tổng.....24
Hình 9.21 Cùng một hợp âm guitar qua hai hiệu ứng.....25
Hình 9.22 Ba phần của một đáp ứng xung, và đường tắt năng lượng của nó.....25
Hình 9.23 Bộ vang Schroeder: bốn lọc lược song song, hai thông toàn phần.....26
Hình 9.24 Biên độ phẳng tuyệt đối, nhưng đáp ứng xung thì không.....27
Hình 9.25 Đo RT60 theo cách T30, và phép kiểm đặt–đo.....28
Hình 9.26 Mật độ tiếng vọng và mốc hoà trộn theo thể tích phòng.....28
Hình 9.27 Tích chập: thay mỗi mẫu bằng một bản sao của đáp ứng xung.....29
Hình 9.28 Chi phí tính: tích chập thẳng so với dùng FFT.....30
Hình 9.29 Vang thuật toán và vang tích chập — sáu điểm so sánh.....31
Hình 9.30 Chương 9 để lại gì cho các chương sau.....34

DANH MỤC BẢNG BIỂU

Bảng 9.1 Sáu mục của chương 9 thuộc họ nào....7
Bảng 9.2 Ba thứ gần giống mà thuộc chương khác....7
Bảng 9.3 Bảng quy đổi Q sang bề rộng dải....18
Bảng 9.4 Bảy bài tập của chương 9....32
Bảng 9.5 Mười ba hàm mới của chương 9....35

CHƯƠNG 9
HIỆU ỨNG PHẦN 2

Mục tiêu của chương

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

- Cài đặt được phép trễ phân số bằng ba kiểu nội suy, và *đo* được cái giá của từng kiểu.
- Giải thích được vì sao flanger dùng nội suy tuyến tính nghe "đục", bằng một con số chứ không bằng cảm nhận.
- Sinh được hệ số của cả bảy kiểu biquad từ ba tham số, và kiểm được tính ổn định của chúng.
- Đọc được đáp ứng biên độ và nói được bộ lọc đang làm gì ở từng vùng tần số.
- Giải thích được vai trò của Q ở cả ba cách nhìn: bề rộng dải, đỉnh nhô tại tần số cắt, và bán kính cực.
- Dựng được một chuỗi EQ nhiều dải và *kiểm lại* được đáp ứng đo được có khớp thiết kế không.
- Nêu được vì sao cắt an toàn hơn nâng, bằng số đo trên cột đỉnh.
- Phân biệt được tremolo và điều chế vòng, và giải thích được vì sao điều chế vòng làm mất cao độ.
- Dựng được bộ vang Schroeder, và *đo* lại RT60 để kiểm xem con số trên giao diện có nghĩa thật hay không.
- Cài đặt được tích chập theo khối và giải thích được vì sao bắt buộc phải dùng FFT.

Một khối mới, sáu mục

Chương 8 làm được rất nhiều việc chỉ bằng phép trễ và phép nhân gain, nhưng nó không chọn được tần số: mọi thứ nó tạo ra đều là răng lược cách đều nhau.

Chương này thêm đúng khối còn thiếu — bộ lọc bậc hai của mục 9.3 — và từ đó ra được cả nửa còn lại của phần hiệu ứng: EQ, wah, và phần đọc đáp ứng của bộ vang.

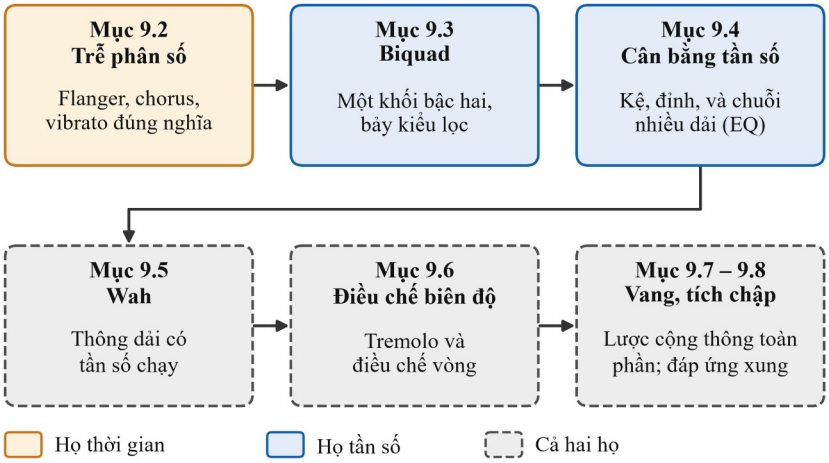
Ba mục có phép kiểm bằng số, và cả ba đều nên chạy trên máy của người đọc: cái giá của nội suy trễ ở mục 9.2, chuỗi EQ ở mục 9.4, và RT60 ở mục 9.7.

9.1 Vị trí của chương trong cuốn sách

9.1.1 Chương 8 để lại một chiều còn thiếu

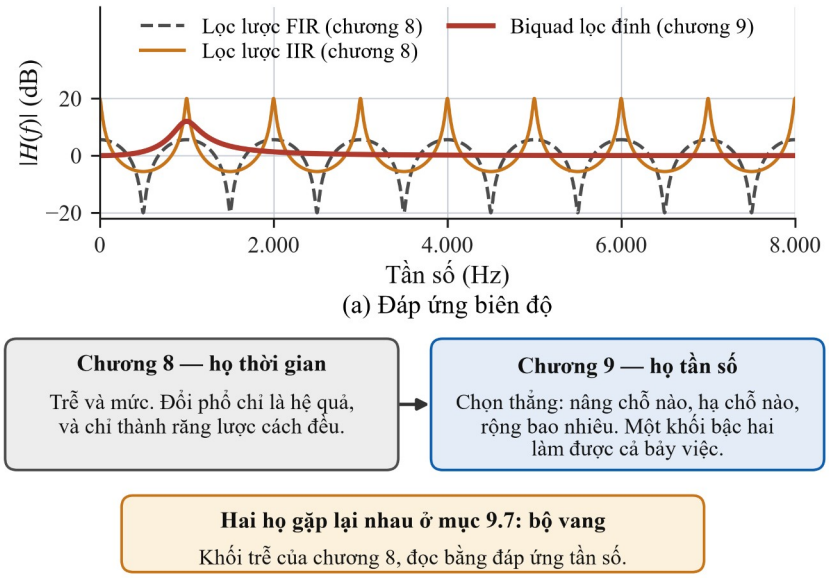
Chương 8 làm được rất nhiều việc chỉ bằng hai phép toán: một phép nhân gain thay đổi theo mức, và một phép cộng với bản trễ của chính nó. Từ hai phép ấy ra được bộ nén, bộ giới hạn, bộ cổng, lọc lược, flanger và tiếng vọng.

Nhưng có một việc chương 8 *không* làm được: chọn tần số. Mọi thứ nó loại ra khỏi phổ đều là răng lược cách đều nhau — muốn hạ riêng dải 300 hertz mà để yên 3 kilohertz thì không có cách nào.



Hình 9.1 Sáu mục kỹ thuật của chương 9

Chương này thêm đúng khối còn thiếu ấy — bộ lọc bậc hai — và từ nó ra được cả nửa còn lại của phần hiệu ứng.



Hình 9.2 Chương 8 chỉ tạo được răng lược; chương 9 chọn được thẳng

9.1.2 Hai họ, và chỗ chúng gặp nhau

Bảng 9.1 Sáu mục của chương 9 thuộc họ nào

Mục	Nội dung	Họ	Dựng trên
9.2	Trễ phân số, flanger, chorus, vibrato	Thời gian	Khối trễ của mục 6.4 và 8.7
9.3	Bộ lọc bậc hai — biquad	Tần số	Hoàn toàn mới
9.4	Cân bằng tần số — EQ	Tần số	Biquad ở 9.3
9.5	Wah — thông dải có tần số chạy	Cả hai	Biquad, nhưng hệ số đổi từng mẫu
9.6	Điều chế biên độ — tremolo, điều chế vòng	Cả hai	Phép nhân, như bộ nén nhưng không đọc mức
9.7 – 9.8	Bộ vang và tích chập	Cả hai	Trễ của chương 8, đọc bằng công cụ tần số của chương 9

Một cách nói gọn về cả chương

Chương 8 hỏi: tín hiệu này *to* bao nhiêu, và tôi trễ nó đi bao lâu.

Chương 9 hỏi một câu khác hẳn: ở tần số này thì tín hiệu bị nâng hay bị hạ, và bao nhiêu.

Cả chương chỉ có một khối mới thật sự — bộ lọc bậc hai của mục 9.3. Năm mục còn lại là cách dùng nó, hoặc cách ghép nó với khối trễ của chương 8.

9.1.3 Những gì chương này không làm

Bảng 9.2 Ba thứ gần giống mà thuộc chương khác

Thứ dễ gộp nhầm	Vì sao không thuộc chương 9	Thuộc chương nào
Panning, âm thanh không gian, HRTF	Chúng dùng trễ và mức để đặt nguồn trong *không gian*. Chương này chỉ dựng một phòng, chưa đặt nguồn vào đâu cả	Chương 10
Nén đa băng	Cần bộ lọc phân băng của chương này ghép với bộ nén của chương 8; và cần cả bài toán bù pha ở chỗ giao băng	Chương 14
Bộ lọc FIR bậc cao, thiết kế cửa sổ	Đó là một chủ đề thiết kế riêng. Chương này dùng đúng một cấu trúc bậc hai, đủ cho mọi việc của phần hiệu ứng	Ngoài phạm vi sách

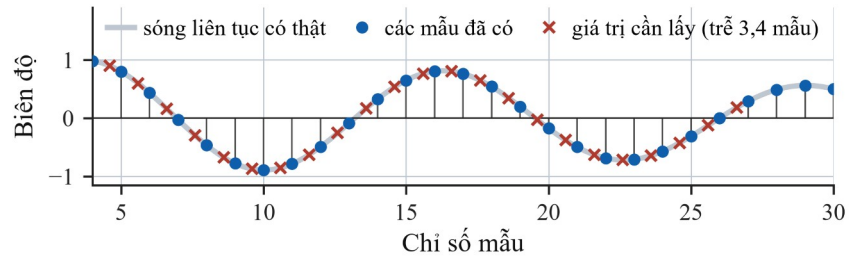
9.1.4 Cách học chương này

- **Ba con số, ba câu hỏi.** Mọi bộ lọc trong chương này chỉ có ba tham số: tần số, lượng nâng, và Q. Chúng trả lời ba câu độc lập nhau — ở đâu, bao nhiêu, rộng bao nhiêu. Nắm được ba câu ấy là nắm được cả mục 9.3 và 9.4.
- **Nhìn đáp ứng, đừng nhìn hệ số.** Năm hệ số của một biquad không nói lên điều gì cho người đọc. Đường |H(f)| thì nói tất cả. Mọi hình trong chương này đều vẽ đáp ứng.
- **Ba chỗ có phép kiểm bằng số.** Nội suy trễ ở mục 9.2, chuỗi EQ ở mục 9.4, và RT60 ở mục 9.7. Cả ba đều đo lại được trên máy của người đọc, và cả ba đều nên đo.

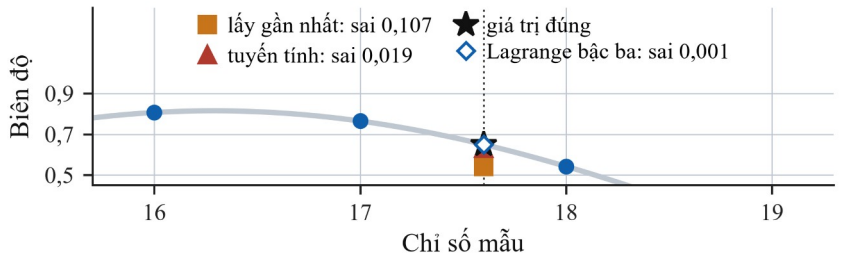
9.2 Trễ phân số và họ điều biến

9.2.1 Vì sao chương 8 chưa đủ

Mục 8.7 đã dựng một bộ trễ điều biến và gọi nó là flanger. Nhưng ở đó có một chỗ được cho qua: độ trễ chạy liên tục thì hầu như không bao giờ rơi đúng vào một số nguyên mẫu.



(a) Trễ 3,4 mẫu: mọi giá trị cần lấy nằm giữa hai mẫu



(b) Ba cách đoán giá trị tại 17,6, và sai số của từng cách

Hình 9.3 Mọi giá trị cần lấy đều nằm giữa hai mẫu đã có

Chương 8 giải quyết bằng một dòng np.interp, tức nội suy tuyến tính. Nó chạy được, nhưng nó không miễn phí — và cái giá của nó là chỗ quyết định một flanger nghe sạch hay nghe đục.

9.2.2 Ba cách đoán giá trị ở giữa

Nội suy tuyến tính nối hai mẫu kề nhau bằng một đường thẳng.

$$y[n] = (1-f)\,x[n-i] + f\,x[n-i-1], \qquad i = \lfloor D \rfloor,\ f = D - i \tag{9.1}$$

D là độ trễ mong muốn tính bằng mẫu, i là phần nguyên, f là phần lẻ. Với f bằng 0 thì công thức trở về phép trễ nguyên của chương 8.

- **Lấy mẫu gần nhất.** Làm tròn D. Rẻ nhất, không có phép nhân nào. Sai tới nửa mẫu, và với hiệu ứng điều biến thì sai số ấy nhảy giật cấp mỗi khi làm tròn đổi — nghe ra là tiếng "rè".
- **Nội suy tuyến tính.** Một phép nhân và một phép cộng. Đủ cho phần lớn công việc. Nhược điểm nêu ở mục 9.2.3.
- **Nội suy lagrange bậc ba.** Đa thức bậc ba qua bốn mẫu liền nhau. Tốn thêm chừng ba phép nhân mỗi mẫu, và sai số nhỏ hơn hẳn.

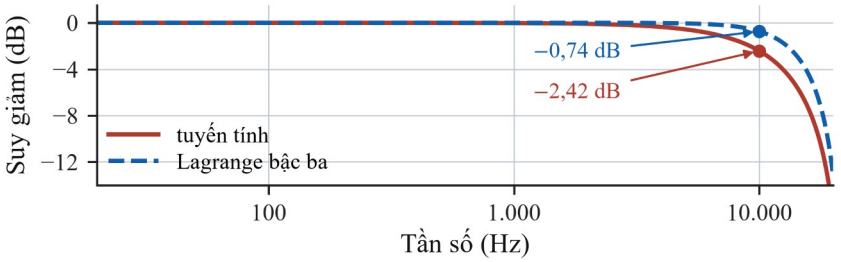
Con số cụ thể ở ô (b) của Hình 9.3: lấy gần nhất sai 0,107; tuyến tính sai 0,019; Lagrange bậc ba sai 0,001. Mỗi bậc tốt hơn bậc trước chừng hai mươi lần.

9.2.3 Cái giá của nội suy tuyến tính — đo được

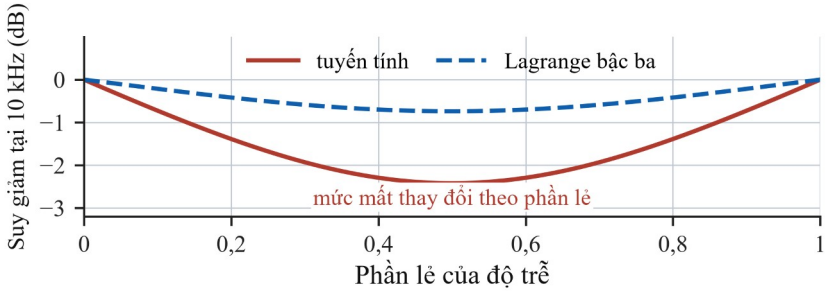
Nội suy tuyến tính không phải là một phép xấp xỉ trung tính: nó là một *bộ lọc thông thấp*. Đáp ứng biên độ của nó tính được thẳng ra từ công thức.

$$|H_{\text{tt}}(\omega)| = \sqrt{1 - 2f(1-f)(1-\cos\omega)} \tag{9.2}$$

Chỉ số dưới tt là "tuyến tính", omega là tần số góc, f là phần lẻ của độ trễ. Khi f bằng 0 hoặc 1 thì biểu thức trong căn bằng 1 và bộ lọc không làm gì; khi f bằng 0,5 thì suy giảm lớn nhất.



(a) Đáp ứng biên độ ở phần lẻ 0,5 (trường hợp xấu nhất)



(b) Mức suy giảm tại 10 kHz khi độ trễ chạy

Hình 9.4 Nội suy tuyến tính mất cao tần, và mức mất ấy THAY ĐỔI

Điều đáng ngại hơn con số tuyệt đối

Ở phần lẻ 0,5, nội suy tuyến tính mất 2,42 decibel tại 10 kilohertz. Lagrange bậc ba chỉ mất 0,74.

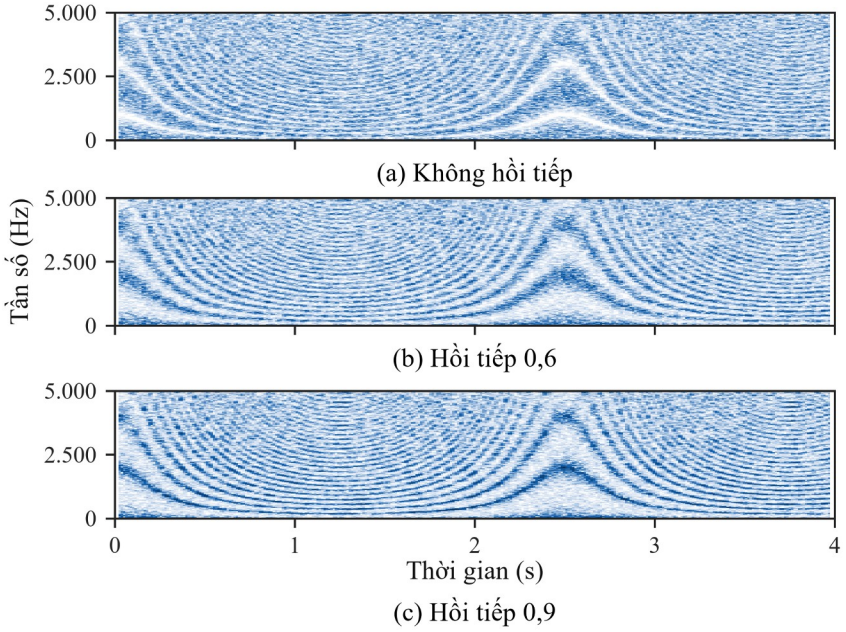
Nhưng điều đáng ngại hơn là con số ấy *không* cố định. Trong một flanger, phần lẻ chạy liên tục từ 0 tới 1 rồi về, nên mức mất cao tần đi từ 0 tới 2,42 decibel rồi về 0 trong mỗi chu kỳ dao động.

Tai nghe ra là tiếng tối dần rồi sáng lại, chồng lên tiếng quét thật của flanger. Kết quả nghe "đục" mà khó chỉ ra vì sao.

Đó là lý do đáng bỏ ba phép nhân mỗi mẫu để dùng Lagrange bậc ba trong mọi hiệu ứng có độ trễ chạy.

9.2.4 Flanger có hồi tiếp

Chương 8 dựng flanger không hồi tiếp: chỉ cộng tín hiệu với bản trễ điều biến của nó. Thêm một đường vòng từ đầu ra về bộ đệm trễ là được flanger đúng nghĩa của nhạc rock.



Hình 9.5 Hồi tiếp làm rãnh sâu hơn, không làm rãnh nhiều hơn

Cần đọc hình này cho đúng. Hồi tiếp *không* tạo thêm rãnh nào — số rãnh vẫn do độ trễ quyết định theo đúng công thức của mục 8.5. Nó làm các rãnh sâu hơn và các đỉnh giữa rãnh nhọn hơn. Đó là toàn bộ khác biệt giữa một flanger nhẹ nhàng và một flanger "như máy bay phản lực".

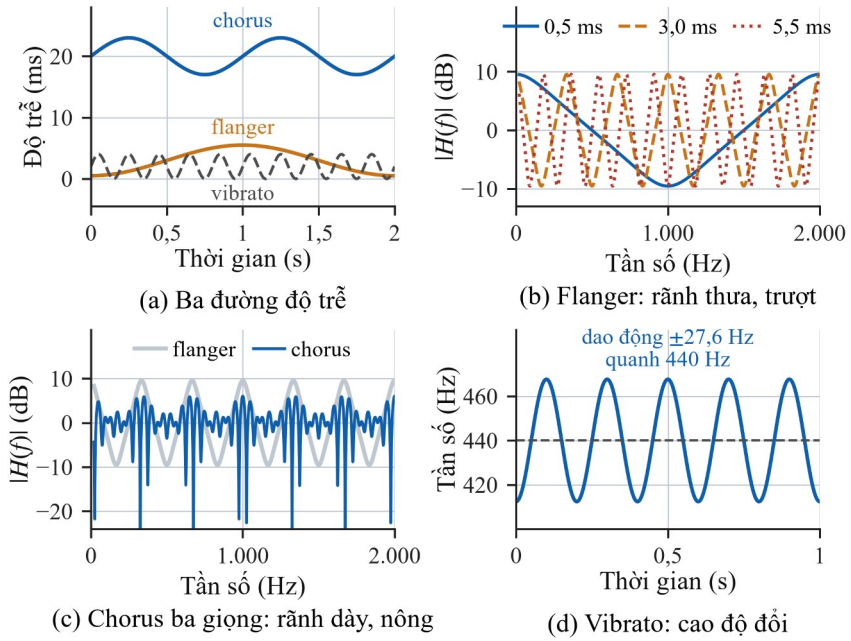
Điều kiện ổn định vẫn nguyên như mục 8.6

Trị tuyệt đối của hệ số hồi tiếp phải nhỏ hơn 1. Không có ngoại lệ nào cho flanger.

Chỗ khác với mục 8.6: bộ đệm trễ ở đây có độ dài *thay đổi*, nên phép nội suy nằm ngay trong vòng hồi tiếp. Sai số nội suy vì thế cũng được nhân lên mỗi vòng.

Hệ quả thực hành: flanger có hồi tiếp cao là chỗ mà nội suy tuyến tính lộ rõ nhất. Nếu buộc phải dùng nội suy tuyến tính thì nên giữ hồi tiếp dưới 0,5.

9.2.5 Ba hiệu ứng, một khối



(a) Ba đường độ trễ

(b) Flanger: rãnh thưa, trượt

(c) Chorus ba giọng: rãnh dày, nông

(d) Vibrato: cao độ đổi

Hình 9.6 Flanger, chorus và vibrato — khác nhau chỉ ở đường độ trễ

Hình (a) là toàn bộ khác biệt giữa ba hiệu ứng: ba đường độ trễ khác nhau, chấm hết. Ba hình còn lại cho thấy hệ quả.

- **Flanger.** Trễ ngắn nên rãnh thưa và nằm trong vùng tai nghe rõ; rãnh *trượt* theo dao động, và đó chính là tiếng quét.
- **Chorus.** Trễ dài hơn và có nhiều giọng lệch pha nhau, nên rãnh dày mà nông. Tai không tách được từng rãnh mà nghe thành "dày" — như nhiều người cùng hát.
- **Vibrato.** *Không* cộng bản gốc, nên không còn lọc lược nào cả. Cái nó đổi là *cao độ*, và phải đo cao độ tức thời mới thấy: với biên độ 2 mili giây ở 5 hertz, cao độ dao động ±27,6 hertz quanh 440 hertz.

Hiệu ứng	Trễ TB	Dao động	Tốc độ	Hồi tiếp	Cộng bản gốc
Flanger	1 – 10 ms	± lớn, gần bằng trễ TB	0,1 – 1 Hz	Có — càng nhiều càng gắt	Có
Chorus	10 – 30 ms	± nhỏ, vài phần trăm	0,5 – 2 Hz	Không	Có
Vibrato	2 – 10 ms	± vừa	4 – 7 Hz	Không	Không

Hai cột "Hồi tiếp" và "Cộng bản gốc" đủ để phân biệt cả ba: chorus có hồi tiếp thì thành flanger; vibrato bỏ bản gốc nên không còn lọc lược, chỉ còn dao động cao độ.

Hình 9.7 Bảng chọn tham số cho ba hiệu ứng điều biến

9.2.6 Cài đặt

Liệt kê 9.1 Nội suy Lagrange bậc ba — bốn hệ số

```
def tre_phan_so_lagrange(x, d):
    """Trễ đi d mẫu, d là SỐ THỰC."""
```

```
    n  = x.size
    i  = np.arange(n) - d
    i0 = np.floor(i).astype(int)
    fr = i - i0                       # phần lẻ, trong [0, 1)

    # bốn hệ số Lagrange qua các mẫu i0-1, i0, i0+1, i0+2
    h = [(-fr * (fr - 1) * (fr - 2)) / 6.0,
         ((fr + 1) * (fr - 1) * (fr - 2)) / 2.0,
         (-(fr + 1) * fr * (fr - 2)) / 2.0,
         ((fr + 1) * fr * (fr - 1)) / 6.0]

    ra = np.zeros(n)
    for k, hk in zip((-1, 0, 1, 2), h):
        j  = i0 + k
        ok = (j >= 0) & (j < n)       # tránh ra ngoài mảng
        ra[ok] += hk[ok] * x[j[ok]]
    return ra
```

Bốn phép kiểm cho phép trễ phân số

Với d nguyên, kết quả phải *trùng* phép trễ thường tới từng bit.

Với d có phần lẻ đúng 0,5 và nội suy tuyến tính, một xung đơn phải chia đôi thành hai mẫu bằng 0,5.

Tổng bốn hệ số Lagrange phải bằng đúng 1 với mọi phần lẻ — nếu không thì tín hiệu bị đổi mức khi độ trễ chạy.

Đo đáp ứng tại 10 kilohertz với phần lẻ 0 và 0,5: phần lẻ 0 phải cho đúng 0 decibel; phần lẻ 0,5 phải cho khoảng −2,4 decibel với tuyến tính và −0,7 với Lagrange bậc ba.

9.3 Bộ lọc bậc hai — biquad

9.3.1 Một khối, bảy việc

Đây là khối mới duy nhất của cả chương, và cũng là khối được dùng lại nhiều nhất trong toàn bộ xử lý âm thanh. Nó có tên là bộ lọc bậc hai, hay biquad — vì hàm truyền của nó là thương của hai tam thức bậc hai.

Phương trình sai phân của nó chỉ có năm phép nhân và bốn phép cộng.

$$y[n] = b_0 x[n] + b_1 x[n-1] + b_2 x[n-2] - a_1 y[n-1] - a_2 y[n-2] \tag{9.3}$$

Ba hệ số b nhân với đầu vào hiện tại và hai đầu vào trước; hai hệ số a nhân với hai đầu ra trước. Dấu trừ trước hai số hạng cuối là quy ước, và giữ đúng quy ước ấy là quan trọng — đảo dấu là bộ lọc thành một thứ khác hẳn.

Viết dưới dạng hàm truyền thì gọn hơn và nói được nhiều hơn.

$$H(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \tag{9.4}$$

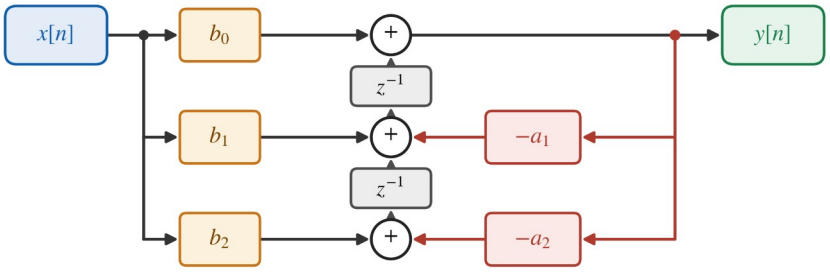
Vì sao bậc hai, không phải bậc một hay bậc bốn

Bậc một chỉ dựng được đường cong đơn điệu — không có đỉnh, không có cộng hưởng, không có Q.

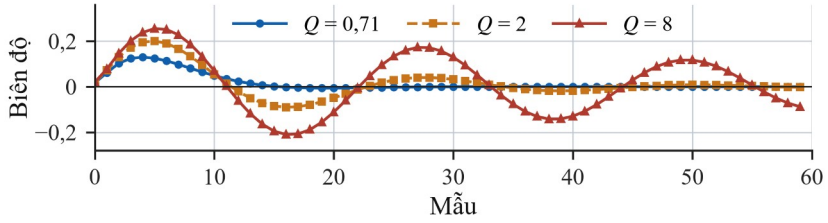
Bậc hai là bậc *thấp nhất* có hai cực liên hợp phức, tức là bậc thấp nhất có cộng hưởng. Từ đó ra đủ bảy kiểu lọc của mục 9.3.4.

Bậc cao hơn thì không cần thiết: nối tiếp nhiều khối bậc hai là ra được bộ lọc bậc bất kỳ, mà lại ổn định số học tốt hơn nhiều so với cài đặt một khối bậc cao. Vì thế mọi thư viện đều làm như thế, kể cả Web Audio ở chương 11.

9.3.2 Cấu trúc cài đặt và hai ô nhớ



(a) Dạng trực tiếp II chuyển vị: chỉ hai ô nhớ



(b) Đáp ứng xung của bộ thông thấp 2 kHz

Hình 9.8 Dạng trực tiếp II chuyển vị, và đáp ứng xung của bộ lọc

Có nhiều cách sắp xếp cùng một phương trình. Dạng dùng trong bài gọi là trực tiếp II *chuyển vị*, và nó được chọn vì hai lý do.

– **Chỉ hai ô nhớ.** Dạng trực tiếp I cần bốn. Hai ô ấy là toàn bộ trạng thái của bộ lọc — chép chúng sang khối sau là chuyển được trạng thái, đúng bài toán của mục 6.9.

– **Ổn định số học tốt hơn.** Với số thực dấu phẩy động 32 bit và bộ lọc Q cao ở tần số thấp, dạng này giữ được độ chính xác tốt hơn hẳn. Với 64 bit thì khác biệt nhỏ, nhưng thói quen nên đúng từ đầu.

Hình (b) cho một điều đáng nhớ: đáp ứng xung của biquad *không* hữu hạn như bộ lọc lược FIR của chương 8 — nó tắt dần mãi, nên gọi là IIR. Q càng cao thì đuôi càng dài và càng dao động.

9.3.3 Ba con số vào, năm hệ số ra

Người dùng không bao giờ nhập năm hệ số. Họ nhập ba con số: tần số trung tâm, Q, và lượng nâng. Việc của mã là đổi ba số ấy thành năm hệ số. Mọi công thức đều đi qua hai đại lượng trung gian.

$$\omega_0 = \frac{2\pi f_c}{f_s}, \qquad \alpha = \frac{\sin\omega_0}{2Q} \tag{9.5}$$

omega không dưới là tần số góc chuẩn hoá, alpha là đại lượng gộp cả Q lẫn tần số. Từ hai số ấy, mỗi kiểu lọc có một công thức riêng. Ví dụ kiểu thông dải — kiểu mà mục 9.5 sẽ dùng cho wah:

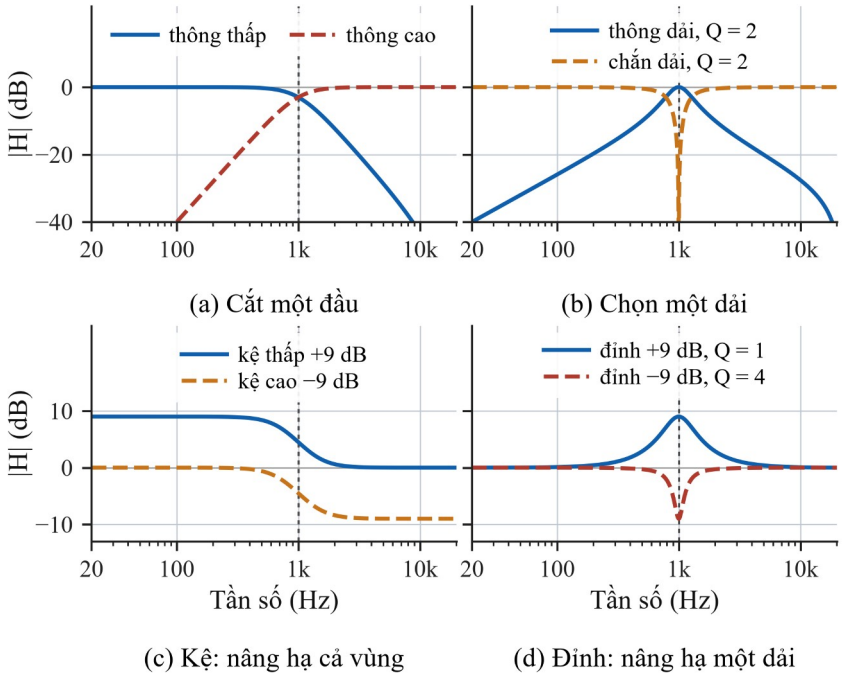
$$b = [\alpha,\ 0,\ -\alpha]\ ; \qquad a = [1+\alpha,\ -2\cos\omega_0,\ 1-\alpha] \tag{9.6}$$

Sau đó chia cả b lẫn a cho a không dưới để chuẩn hoá về dạng của công thức (9.3). Sáu kiểu còn lại chỉ khác ở ba dòng tính b và a; toàn bộ nằm trong sổ tay RBJ Audio EQ Cookbook, và mô-đun mã nguồn của sách chép đúng theo đó.

9.3.4 Bảy kiểu lọc

TÊN HÀM	TÊN VIỆT	LÀM GÌ / DÙNG ĐỂ
`thong_thap`	Thông thấp	Cắt trên fc, dốc 12 dB/quãng tám *Loại tiếng rít, làm dịu nguồn*
`thong_cao`	Thông cao	Cắt dưới fc, dốc 12 dB/quãng tám *Loại ù nguồn điện, tiếng gió micro*
`thong_dai`	Thông dải	Chỉ giữ quanh fc, bề rộng do Q *Bộ lọc của wah; tách một dải để phân tích*
`chan_dai`	Chắn dải	Tạo một rãnh hẹp tại fc *Loại ù 50 Hz và các hài của nó*
`ke_thap`	Kệ thấp	Nâng hoặc hạ CẢ vùng dưới fc *Chỉnh lượng trầm của cả bản*
`ke_cao`	Kệ cao	Nâng hoặc hạ CẢ vùng trên fc *Chỉnh độ sáng; thước K (chương 7)*
`dinh`	Đỉnh (peaking)	Nâng hoặc hạ MỘT dải quanh fc *Dải giữa của EQ đồ hoạ và bán tham số*

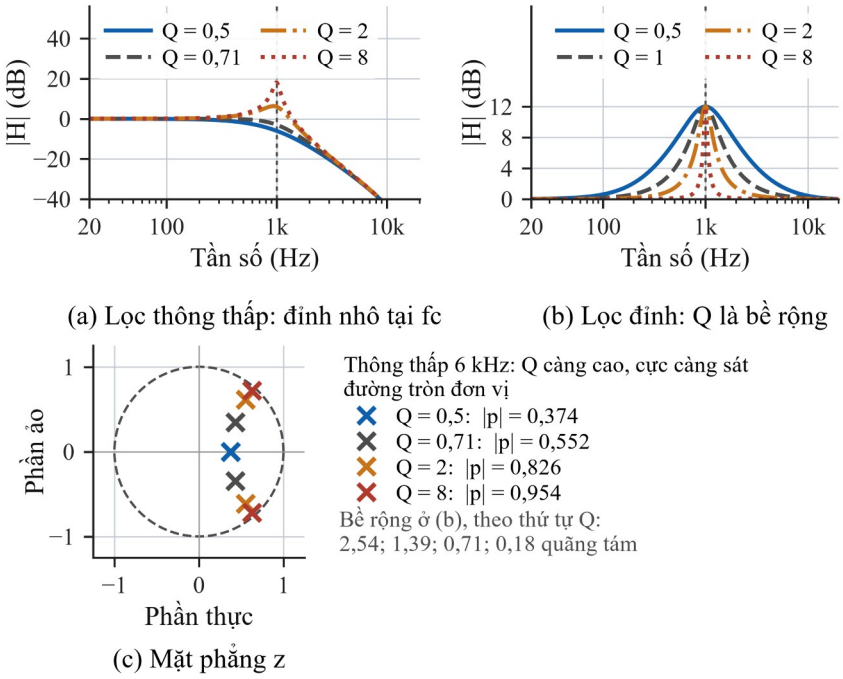
Hình 9.9 Bảy kiểu biquad và việc của từng kiểu



Hình 9.10 Bảy kiểu chia thành bốn việc, nhìn trên đáp ứng biên độ

Bảy kiểu chia thành bốn việc rõ rệt: cắt hẳn một đầu, chọn hoặc bỏ một dải hẹp, nâng hạ cả một vùng, nâng hạ đúng một dải. Người dùng EQ hằng ngày chỉ cần ba kiểu cuối, nhưng cả bảy đều là cùng một khối với năm hệ số khác nhau.

9.3.5 Q — con số hay bị hiểu sai nhất



Hình 9.11 Ba cách nhìn cùng một con số Q

Q trả lời câu hỏi "hẹp bao nhiêu". Nhưng với bộ lọc *cắt*, nó còn làm một việc nữa mà nhiều người không ngờ: nó dựng một bướu ngay tại tần số cắt. Với Q bằng 8, bướu ấy cao 18 decibel.

Ba giá trị Q đáng thuộc

Q = 0,7071 (một phần căn hai) là ngưỡng *không* có bướu — đáp ứng phẳng nhất có thể ở dải thông. Đây là giá trị mặc định của mọi bộ lọc cắt, và là giá trị nên dùng khi chỉ muốn cắt.

Q = 0,5 cho hai cực *thực* — bộ lọc không còn cộng hưởng chút nào, đổi lại chuyển tiếp mềm hơn.

Q = 1,4142 (căn hai) với *lọc đỉnh* đúng bằng bề rộng một quãng tám. Đây là mốc để đổi qua lại giữa hai cách nghĩ.

Ô (c) của Hình 9.11 là cách nhìn chặt nhất: hai cực của bộ lọc nằm càng sát đường tròn đơn vị thì Q càng cao. Với thông thấp 6 kilohertz, bán kính cực đi từ 0,374 ở Q bằng 0,5 lên 0,954 ở Q bằng 8.

Cực ra ngoài đường tròn là mất ổn định

Đây đúng cùng một điều kiện với vòng hồi tiếp ở mục 8.6, chỉ phát biểu bằng ngôn ngữ khác. Ở đó là |g| < 1; ở đây là mọi cực phải nằm trong đường tròn đơn vị.

Với biquad tính đúng theo công thức RBJ, điều kiện này luôn được thoả. Nó chỉ không còn đúng khi tự chỉnh tay hệ số, hoặc khi Q rất cao và tần số rất thấp làm sai số dấu phẩy động đẩy cực ra ngoài.

Hàm on_dinh_biquad của mô-đun chỉ có một dòng và nên gọi mỗi khi sinh hệ số từ tham số người dùng nhập vào.

9.3.6 Cài đặt

Liệt kê 9.2 Biquad — sinh hệ số và chạy bộ lọc

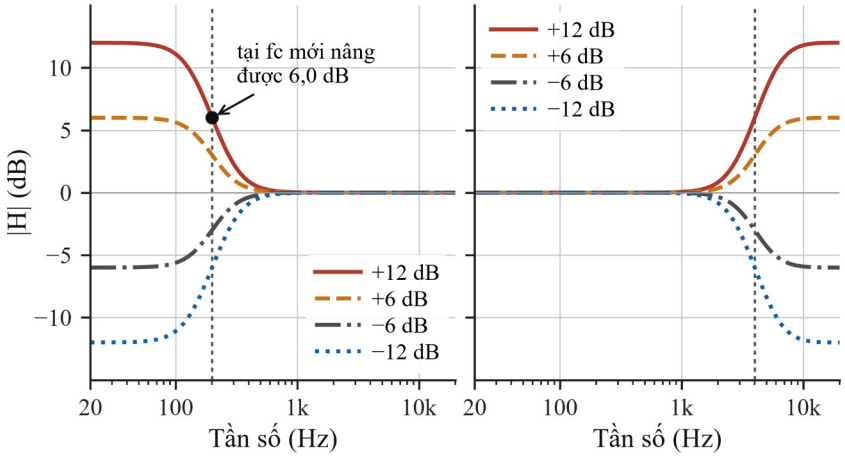
```
def biquad_thong_dai(fc, Q, sr=44100):
    w0 = 2.0 * math.pi * fc / sr
    cw, sw = math.cos(w0), math.sin(w0)
    al = sw / (2.0 * Q)
    b = [al, 0.0, -al]
    a = [1 + al, -2 * cw, 1 - al]
    return np.array(b) / a[0], np.array(a) / a[0]


def ap_biquad(x, b, a):
    """Dạng trực tiếp II chuyển vị – chỉ hai biến trạng thái."""
    y = np.empty(x.size)
    z1 = z2 = 0.0                    # TOÀN BỘ bộ nhớ
    for n in range(x.size):
        o  = b[0] * x[n] + z1
        z1 = b[1] * x[n] - a[1] * o + z2
        z2 = b[2] * x[n] - a[2] * o
        y[n] = o
    return y
```

9.4 Cân bằng tần số

9.4.1 Kệ — kiểu dùng nhiều nhất mà ít được giải thích nhất

Bộ lọc kệ nâng hoặc hạ *cả* một vùng chứ không phải một dải. Vì thế nó đổi cân bằng chung của bản mà không tạo ra chỗ nhô nào — đúng thứ cần khi bản dựng "thiếu trầm" hoặc "chói quá".



(a) Kệ thấp, fc = 200 Hz (b) Kệ cao, fc = 4 kHz

Hình 9.12 Kệ thấp và kệ cao, bốn lượng nâng mỗi loại

Điều cần nhớ về tần số của kệ

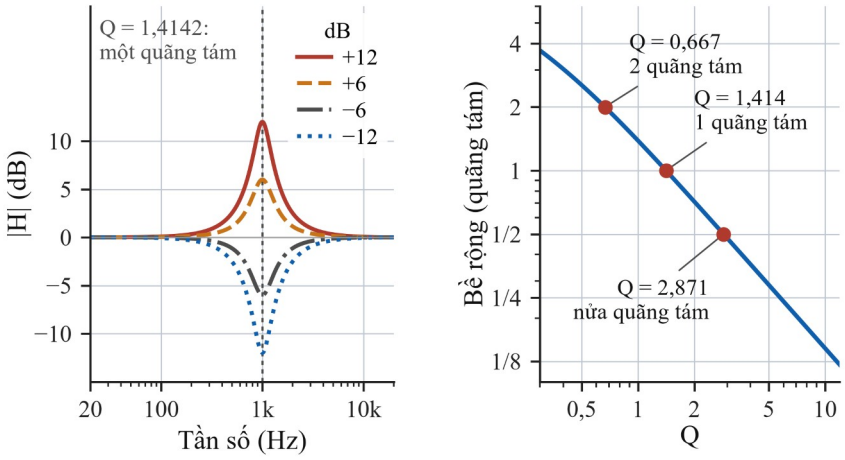
Tại đúng fc, kệ mới đi được *nửa* lượng nâng. Với kệ thấp +12 decibel đặt ở 200 hertz, tại 200 hertz mới nâng được 6,0 decibel.

Hệ quả thực hành: muốn nâng đủ 12 decibel ở 100 hertz thì phải đặt fc cao hơn 100 hertz khá nhiều — chừng 250 tới 300 hertz.

Đây là chỗ khác hẳn lọc đỉnh: lọc đỉnh cho đúng lượng nâng tại đúng fc.

Thước K của chương 7 chính là một bộ lọc kệ cao khoảng +4 decibel nối tiếp với một bộ thông cao — nghĩa là mục 7.2 đã dùng đúng khối này, chỉ chưa gọi tên.

9.4.2 Lọc đỉnh và bề rộng dải



(a) Lọc đỉnh tại 1 kHz (b) Đổi Q sang quãng tám

Hình 9.13 Lọc đỉnh, và cách đổi Q sang quãng tám

Lọc đỉnh là dải giữa của mọi bộ EQ. Ba con số của nó trả lời ba câu độc lập nhau: tần số nói "ở đâu", lượng nâng nói "bao nhiêu", Q nói "rộng bao nhiêu".

Nhưng người dùng không nghĩ bằng Q — họ nghĩ bằng quãng tám. Hai công thức sau là cầu nối, và chúng đi được cả hai chiều.

$$BW = \frac{2}{\ln 2}\operatorname{arcsinh}\frac{1}{2Q}, \qquad Q = \frac{1}{2\sinh\left(\frac{\ln 2}{2}BW\right)} \tag{9.7}$$

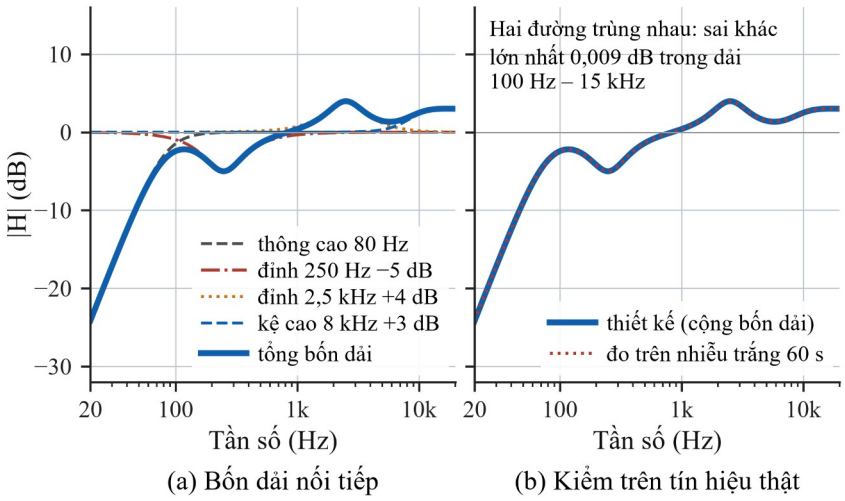
Bảng 9.3 Bảng quy đổi Q sang bề rộng dải

Q	Bề rộng (quãng tám)	Dùng khi
0,667	2,00	Chỉnh cân bằng chung, nghe tự nhiên
1,414	1,00	Mốc chuẩn — dải giữa của EQ đồ hoạ
2,871	0,50	Nhấn hoặc hạ một nhạc cụ trong bản trộn
4,318	0,33	Sửa một chỗ cụ thể mà không đụng xung quanh
8,650	0,17	Diệt một tần số cộng hưởng của phòng

9.4.3 Chuỗi nhiều dải — decibel thì cộng lại

Một bộ EQ nhiều dải chỉ là nhiều biquad nối tiếp nhau. Nối tiếp là *nhân* các đáp ứng với nhau, mà nhân trong thang tuyến tính chính là *cộng* trong thang decibel.

$$H_{\text{tổng}}(f) = \prod_k H_k(f) \Longrightarrow L_{\text{tổng}}(f) = \sum_k L_k(f) \ \text{[dB]} \tag{9.8}$$



(a) Bốn dải nối tiếp (b) Kiểm trên tín hiệu thật

Hình 9.14 Bốn dải nối tiếp, và phép kiểm trên tín hiệu thật

Ba hệ quả của việc decibel cộng lại

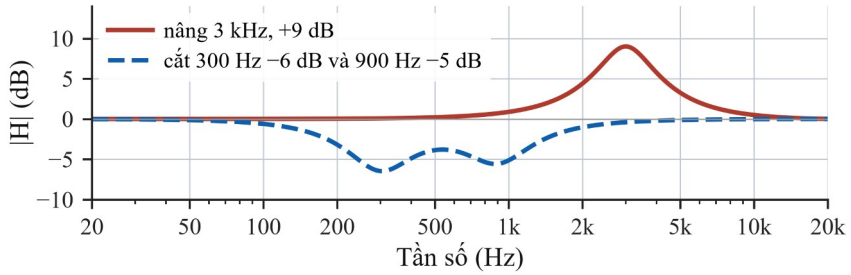
Đường tổng đọc được bằng mắt: cứ cộng các đường thành phần lại. Không cần tính gì cả.

Thứ tự các dải *không* quan trọng về mặt biên độ — nhân thì giao hoán. (Về pha thì cũng vậy, vì pha cũng cộng.)

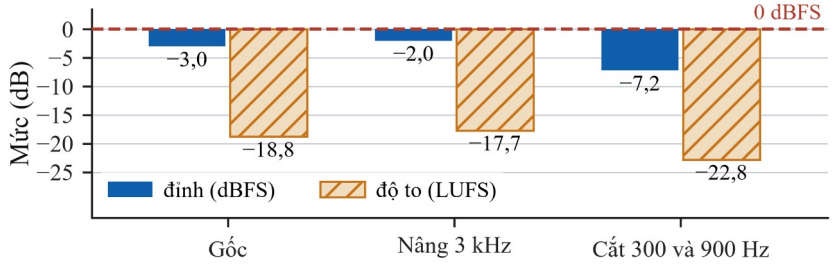
Một bộ EQ mười dải không tốn gì hơn mười lần một khối bậc hai. Đó là lý do EQ là hiệu ứng rẻ nhất trong mọi hiệu ứng.

Hình (b) là phép kiểm nên làm với mọi chuỗi EQ tự viết: cho nhiễu trắng đi qua, đo phổ trước và sau, lấy hiệu, rồi so với đường thiết kế. Trong bài, hai đường trùng nhau tới dưới một phần trăm decibel trong dải 100 hertz tới 15 kilohertz.

9.4.4 Cắt hay nâng — một quy tắc có căn cứ bằng số



(a) Hai cách làm cho tiếng sáng hơn



(b) Cái giá ở cột đỉnh

Hình 9.15 Hai cách làm tiếng sáng hơn, và cái giá ở cột đỉnh

Hai đường ở hình (a) nghe ra gần giống nhau: cả hai đều làm tăng tỉ lệ năng lượng ở vùng cao. Nhưng chúng khác hẳn nhau ở một chỗ đo được.

- **Nâng thì đẩy đỉnh lên.** Nâng 3 kilohertz thêm 9 decibel đưa đỉnh từ −3,0 lên −2,0 dBFS. Ăn mất headroom, và nếu bản đã gần trần thì cắt méo ngay.
- **Cắt thì hạ đỉnh xuống.** Cắt 300 và 900 hertz đưa đỉnh xuống −7,2 dBFS. Nghe nhỏ hơn, nhưng chỉ cần bù lại bằng một hằng số gain là xong — mà hằng số ấy không đổi hình dạng phổ.

Quy tắc lâu đời của phòng thu

Cắt thì rộng tay, nâng thì dè dặt.

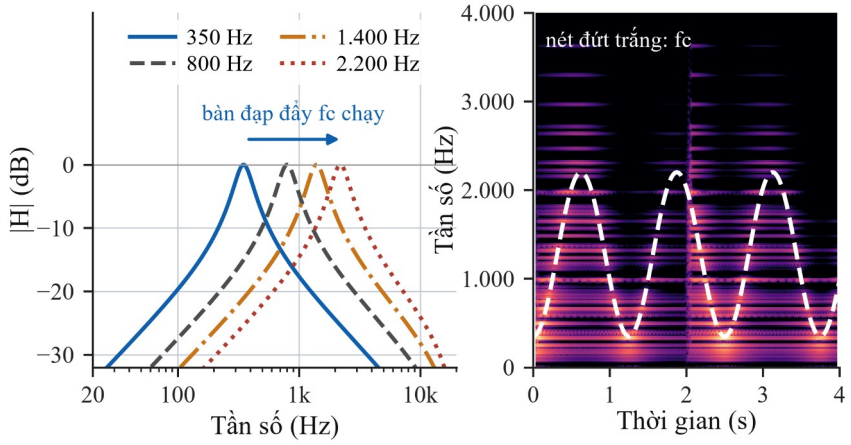
Lý do kỹ thuật là cột đỉnh ở hình (b). Còn một lý do nữa thuộc về tai: khi muốn nghe rõ một nhạc cụ, hạ những nhạc cụ *che* nó thường hiệu quả hơn là nâng chính nó — vì nâng thì làm dày thêm chỗ vốn đã chật.

Ngoại lệ: kệ cao vài decibel để thêm "không khí" cho cả bản là việc nâng hợp lý và rất hay dùng.

9.5 Wah — bộ lọc có tần số chạy

9.5.1 Không có khối nào mới

Wah là bộ thông dải của mục 9.3 với đúng một thay đổi: hệ số được tính lại ở *từng* mẫu, vì tần số trung tâm đổi liên tục theo bàn đạp.



(a) Thông dải có fc chạy (b) Trên hợp âm guitar

Hình 9.16 Wah là một bộ thông dải có tần số trung tâm chạy

Trên phổ đồ, vệt sáng đi theo đúng đường bàn đạp — đó chính là cái tai nghe ra thành tiếng "oa-oa". Dải quét thường từ chừng 350 tới 2.200 hertz, đúng vùng mà tai nhạy nhất và cũng đúng vùng của các nguyên âm trong tiếng nói.

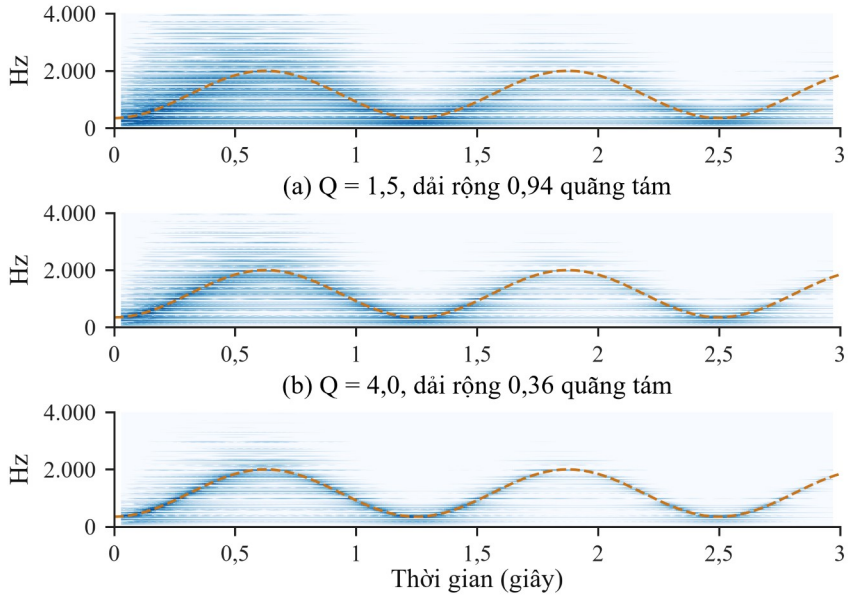
Vì sao wah nghe như tiếng nói

Bộ máy phát âm của người tạo ra các đỉnh cộng hưởng gọi là formant, và vị trí của chúng quyết định chúng ta nghe ra nguyên âm nào.

Formant thứ nhất của các nguyên âm nằm trong khoảng 300 – 800 hertz, formant thứ hai trong khoảng 900 – 2.500 hertz. Dải quét của wah phủ đúng vùng ấy.

Vì thế khi bàn đạp đi lên rồi xuống, tai nghe ra một chuỗi gần giống nguyên âm "oa" — và cái tên của hiệu ứng chính là mô phỏng âm thanh đó.

9.5.2 Q quyết định wah nghe mềm hay gắt



(a) Q = 1,5, dải rộng 0,94 quãng tám

(b) Q = 4,0, dải rộng 0,36 quãng tám

(c) Q = 12,0, dải rộng 0,12 quãng tám

Hình 9.17 Ba giá trị Q trên cùng một đoạn guitar

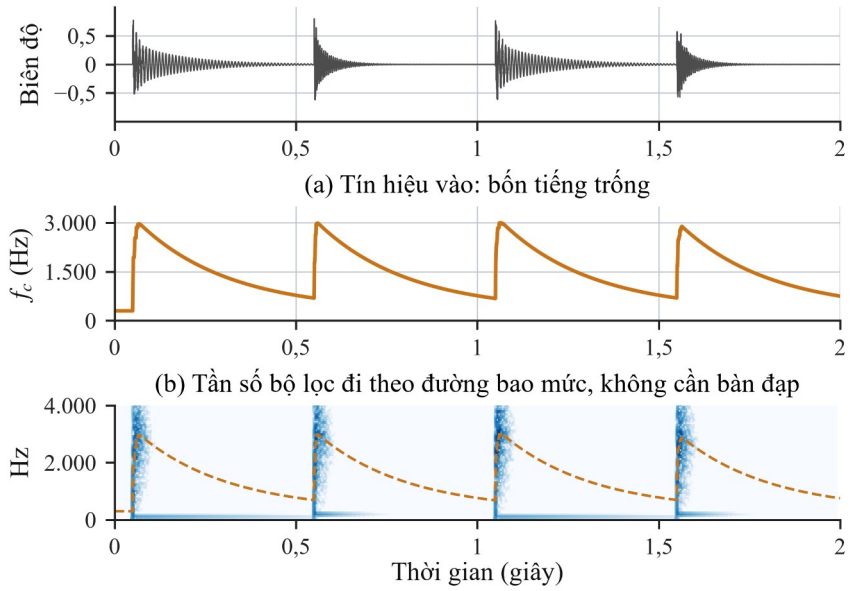
Q thấp thì dải rộng, hiệu ứng nhẹ và giữ được nhiều thân tiếng. Q cao thì dải rất hẹp, tiếng thành mảnh và rõ chất tổng hợp. Khoảng thường dùng là 2 tới 6; trên 10 thì bộ lọc gần mất ổn định và tiếng bắt đầu rít.

Hai chỗ dễ sai khi tự viết wah

Tính lại hệ số ở từng mẫu là đúng, nhưng *không* được đặt lại hai biến trạng thái z1 và z2 khi tính lại. Đặt lại là mất hết bộ nhớ của bộ lọc và tiếng ra thành một chuỗi gai.

Khi fc đổi rất nhanh — bàn đạp giật hoặc auto-wah trên tiếng gõ sắc — bộ lọc có thể sinh gai ngay tại chỗ đổi. Cách chữa: làm trơn đường fc bằng một bộ lọc một cực trước khi đưa vào, đúng khối duong_bao_muc của mục 8.2.

9.5.3 Auto-wah — nối hai họ hiệu ứng lại



Hình 9.18 Auto-wah: tần số bộ lọc đi theo đường bao mức

Auto-wah dùng lại nguyên khối bộ dò mức của mục 8.2: chơi mạnh thì bộ lọc mở lên cao, chơi nhẹ thì nó đóng xuống. Không có bàn đạp nào cả, mà hiệu ứng vẫn đi đúng theo nhịp — vì nhịp nằm sẵn trong đường bao.

Đây là ví dụ rõ nhất trong cả cuốn sách cho việc hai họ hiệu ứng ghép được vào nhau: một khối của chương 8 điều khiển một khối của chương 9. Cùng ý tưởng ấy còn cho ra bộ nén theo đường bên và bộ lọc theo đường bao ở chương 14.

Liệt kê 9.3 Auto-wah — chỉ khác wah thường ở ba dòng

```
def auto_wah(x, f_dau=300.0, f_cuoi=3000.0, Q=4.0,
             tan_cong_ms=5.0, nha_ms=120.0, sr=44100):
    # KHỐI CỦA CHƯƠNG 8: đường bao mức
    env = duong_bao_muc(x, tan_cong_ms, nha_ms, sr, "dinh")
    e   = env / max(env.max(), 1e-12)

    # KHỐI CỦA CHƯƠNG 9: thông dải, fc đổi từng mẫu
    fc = f_dau + (f_cuoi - f_dau) * e
    return thong_dai_quet(x, fc, Q, sr)
```

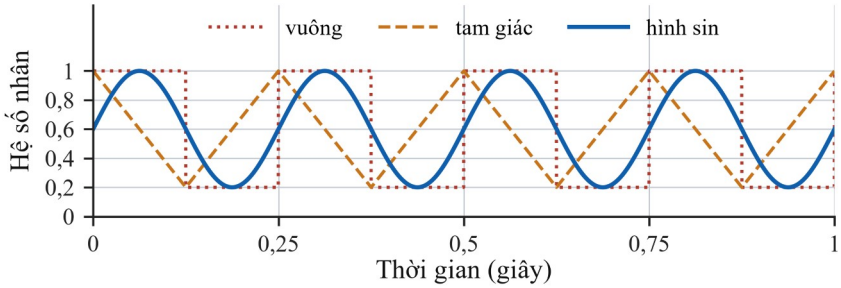
9.6 Điều chế biên độ

9.6.1 Tremolo — phép nhân đơn giản nhất của cả cuốn sách

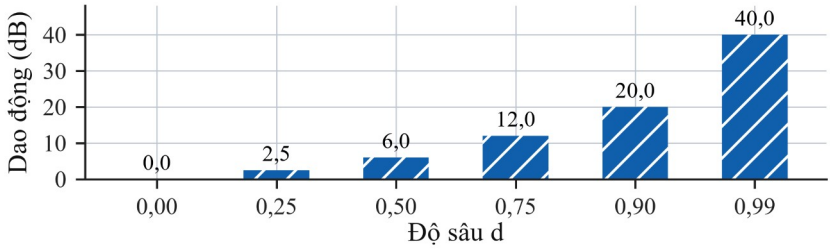
Tremolo là nhân tín hiệu với một dao động *chậm* và *luôn không âm*. Hai điều kiện ấy là toàn bộ định nghĩa, và chúng phân biệt tremolo với điều chế vòng ở mục sau.

$$y[n] = x[n]\left[1 - d + d \cdot \frac{1 + \sin(2\pi f_{\text{lfo}} n / f_s)}{2}\right] \tag{9.9}$$

d là độ sâu, từ 0 tới 1; f dưới chỉ số lfo là tốc độ dao động, thường 2 tới 8 hertz. Phần trong ngoặc vuông luôn nằm giữa 1 − d và 1, nên nó không bao giờ âm và không bao giờ vượt 1.



(a) Ba dạng dao động, độ sâu 0,8 và tốc độ 4 Hz



(b) Độ sâu (thang tuyến tính) đổi sang decibel

Hình 9.19 Ba dạng dao động, và cách độ sâu đổi sang decibel

Một chỗ hay gây nhầm lẫn khi tự viết giao diện

Độ sâu là thang *tuyến tính*, còn tai nghe theo thang decibel. Độ sâu 0,25 chỉ cho 2,5 decibel dao động; độ sâu 0,9 cho 20; độ sâu 0,99 cho 40.

Hệ quả: một núm chạy tuyến tính từ 0 tới 1 sẽ có nửa đầu gần như không nghe ra gì, còn năm phần trăm cuối thì đổi rất dữ dội.

Cách chữa: cho núm chạy theo *decibel* rồi đổi ngược ra độ sâu, hoặc dùng thang bình phương. Cùng vấn đề với núm "feedback" của mục 8.6 — và cùng một cách chữa.

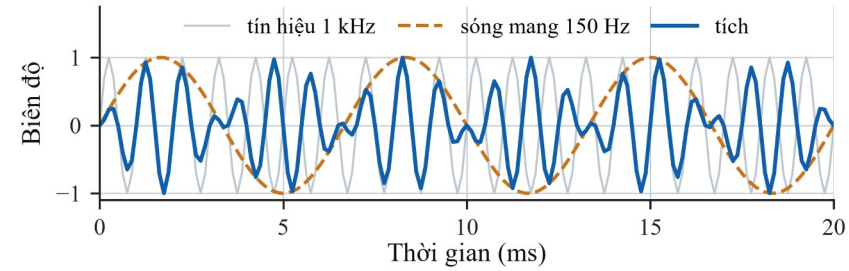
9.6.2 Điều chế vòng — đổi một dấu, đổi cả kết quả

Điều chế vòng cũng là một phép nhân, khác tremolo đúng hai chỗ: tần số của sóng mang nằm trong vùng *nghe được*, và hệ số nhân *đổi dấu*.

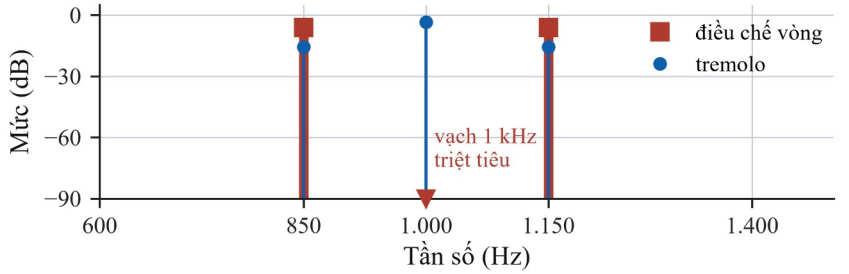
Vì sao hai chỗ ấy đổi hẳn kết quả thì một công thức lượng giác phổ thông trả lời gọn hơn mọi lời giải thích.

$$\sin(2\pi f_t t)\sin(2\pi f_m t) = \frac{1}{2}\cos(2\pi(f_t - f_m)t) - \frac{1}{2}\cos(2\pi(f_t + f_m)t) \qquad (9.10)$$

Chỉ số dưới t là tín hiệu, m là sóng mang. Tích của hai hình sin *không* còn chứa tần số nào của hai hình sin ấy — nó chỉ chứa hiệu và tổng của chúng.



(a) Nhân với sóng mang có đổi dấu



(b) Phổ: tremolo giữ vạch gốc, điều chế vòng thì không

Hình 9.20 Sóng mang biến mất; chỉ còn hai dải bên ở hiệu và tổng

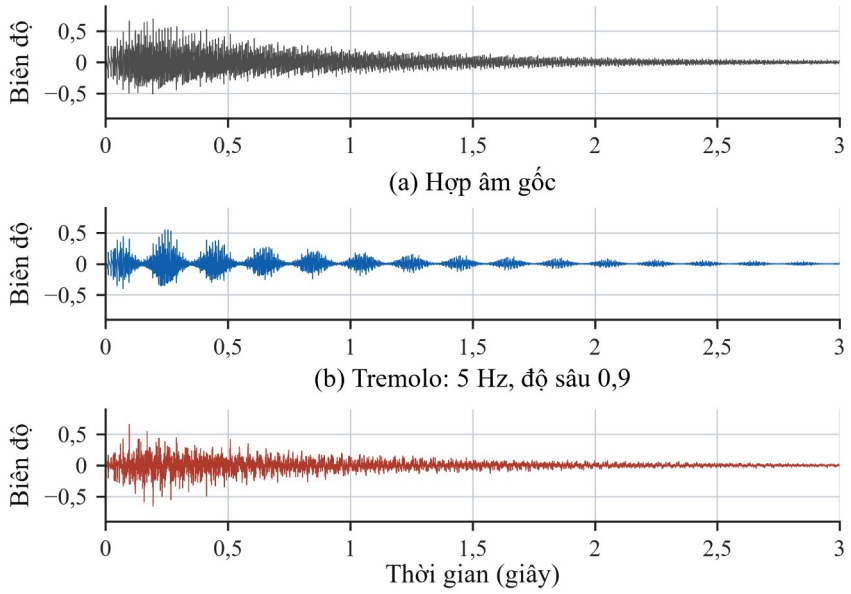
Hình (b) là phép kiểm trực tiếp: tín hiệu 1 kilohertz nhân với sóng mang 150 hertz cho đúng hai vạch ở 850 và 1.150 hertz, còn vạch 1 kilohertz triệt tiêu hoàn toàn. Tremolo ở cùng 150 hertz thì giữ nguyên vạch gốc — vì hệ số của nó không âm nên trong nó còn một thành phần một chiều.

Vì sao điều chế vòng nghe kim loại"

Một nhạc cụ có cao độ rõ vì các thành phần của nó là *bội số* nguyên của tần số cơ bản — 100, 200, 300, 400 hertz.

Điều chế vòng với sóng mang 60 hertz biến chúng thành 40, 160, 140, 260, 240, 360... Không còn bội số nguyên của bất kỳ số nào, nên tai không dựng được cao độ.

Đó chính là cái chúng ta gọi là "kim loại" hay "như người máy". Và đó cũng là lý do điều chế vòng gần như chỉ dùng làm hiệu ứng đặc biệt, chứ không dùng trên cả bản nhạc.



(a) Hợp âm gốc

(b) Tremolo: 5 Hz, độ sâu 0,9

(c) Điều chế vòng: sóng mang 220 Hz

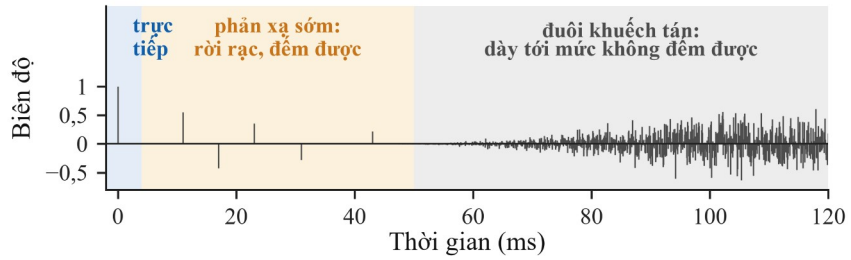
Hình 9.21 Cùng một hợp âm guitar qua hai hiệu ứng

Nhìn dạng sóng là phân biệt được ngay. Tremolo giữ nguyên hình dạng bên trong và chỉ nắn đường bao. Điều chế vòng đổi hẳn nội dung bên trong.

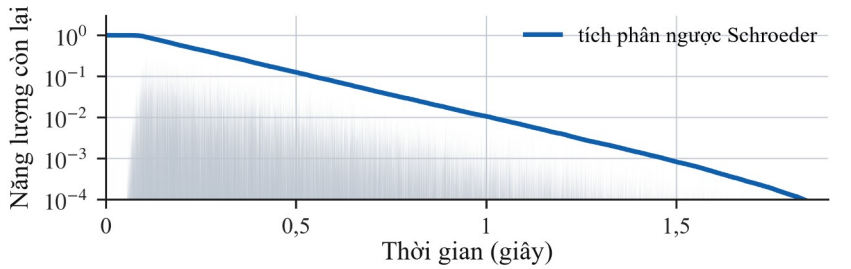
9.7 Bộ vang

9.7.1 Một đáp ứng xung phòng gồm ba phần

Trước khi dựng một bộ vang, phải biết đang dựng cái gì. Câu trả lời gọn nhất: một *đáp ứng xung* phòng — thứ chúng ta nghe được khi vỗ tay một cái trong phòng ấy.



(a) Ba phần của đáp ứng xung (120 ms đầu)



(b) Cả đáp ứng xung và đường tắt dần năng lượng (thang loga)

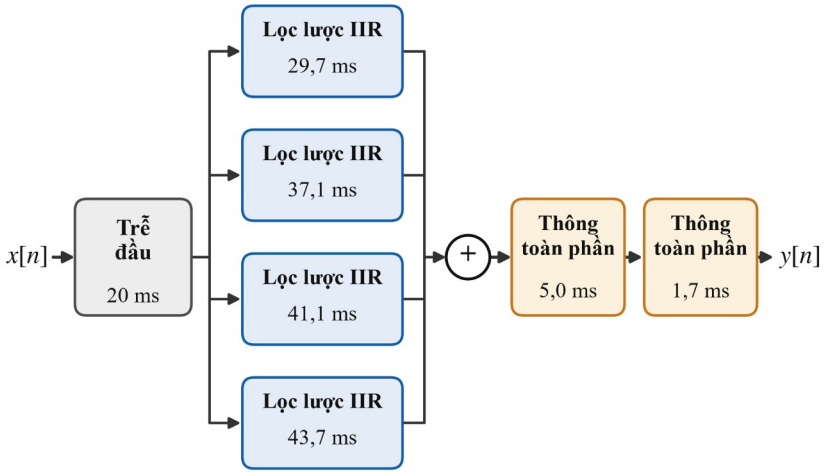
Hình 9.22 Ba phần của một đáp ứng xung, và đường tắt năng lượng của nó

– **Tiếng trực tiếp.** Đến trước nhất, đi thẳng từ nguồn tới tai. Nó cho biết nguồn ở đâu.

– **Phản xạ sớm.** Vài chục mili giây đầu, rời rạc và *đếm được*. Chúng cho biết phòng to cỡ nào và có hình gì — vì khoảng cách giữa chúng chính là hiệu quãng đường.
– **Đuôi vang khuếch tán.** Sau mốc hoà trộn, dày tới mức không đếm được từng tiếng. Nó cho biết phòng hút âm nhiều hay ít, tức là RT60 bao nhiêu.

Một bộ vang thuật toán phải dựng đủ cả ba, và mục này dựng từng phần một.

9.7.2 Sơ đồ Schroeder



Bốn độ trễ nguyên tố cùng nhau để các đỉnh cộng hưởng không chồng lên nhau.
Bốn lọc lược song song dựng đuôi vang; hai thông toàn phần nối tiếp làm dày mật độ.

Hình 9.23 Bộ vang Schroeder: bốn lọc lược song song, hai thông toàn phần

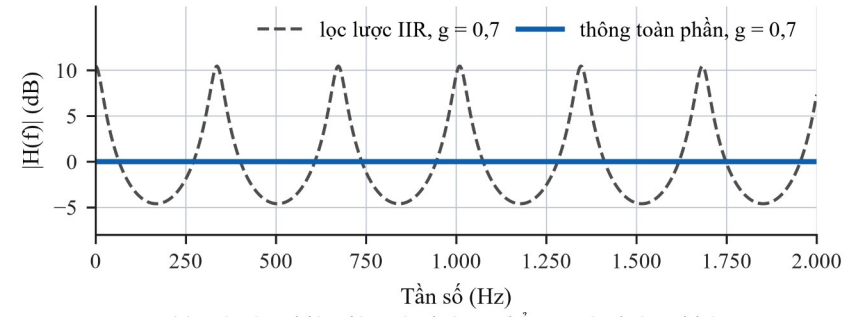
Sơ đồ này có từ năm 1962 và vẫn là nền tảng của mọi bộ vang thuật toán. Nó gồm ba tầng.

– **Trễ đầu.** Một phép trễ thường, chừng 10 tới 40 mili giây. Nó tách tiếng trực tiếp khỏi đuôi vang, và tai đọc nó thành khoảng cách tới bức tường gần nhất.
– **Bốn bộ lọc lược IIR song song.** Chúng dựng đuôi vang và quyết định RT60. Bốn độ trễ phải chọn nguyên tố cùng nhau — nếu không, các đỉnh cộng hưởng của chúng chồng lên nhau và tai nghe ra một tiếng ngân giả rất khó nghe.
– **Hai bộ thông toàn phần nối tiếp.** Chúng làm dày mật độ tiếng vọng. Lý do phải là thông toàn phần chứ không phải lọc lược nằm ở mục sau.

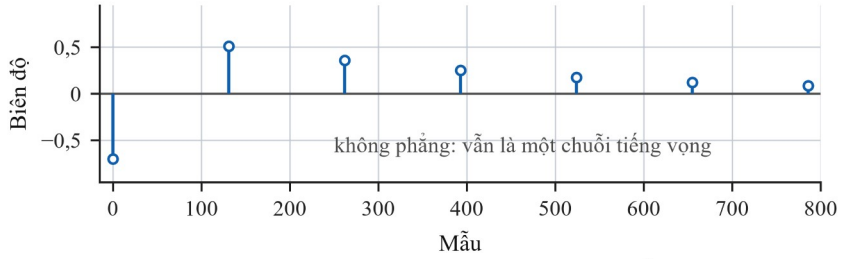
9.7.3 Bộ thông toàn phần — khối lạ nhất của cuốn sách

$$y[n] = -g\,x[n] + x[n-D] + g\,y[n-D]\,; \qquad |H(f)| = 1 \;\; \forall f \tag{9.11}$$

Biên độ bằng đúng 1 ở *mọi* tần số. Nó không nâng, không hạ, không loại bỏ gì cả — nó chỉ đổi *pha*.



(a) Đáp ứng biên độ: một đường phẳng, một đường không



(b) Đáp ứng xung của bộ thông toàn phần

Hình 9.24 Biên độ phẳng tuyệt đối, nhưng đáp ứng xung thì không

Vì sao schroeder cần đúng khối này

Bốn bộ lọc lược đã tạo đủ rãnh vào phổ rồi. Thêm bất kỳ khối nào có đáp ứng không phẳng nữa là phổ thành lỗ chỗ và tiếng vang nghe "màu mè".

Nhưng vẫn cần làm dày mật độ tiếng vọng, vì bốn bộ lược đứng riêng chỉ cho bốn chuỗi vọng đều đặn — tai đếm được và nghe ra tiếng "lạch cạch".

Bộ thông toàn phần giải đúng bài toán ấy: nó tán các tiếng vọng ra thành nhiều tiếng nhỏ hơn mà không đụng một decibel nào vào phổ. Nghe riêng nó thì gần như không nghe ra gì; đặt sau bốn bộ lược thì đổi hẳn chất lượng.

9.7.4 Đo RT60 — con số duy nhất phải đúng

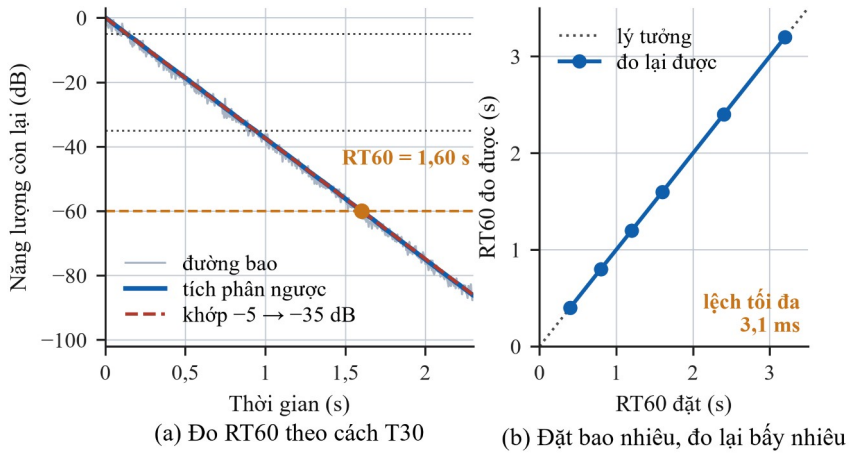
Đường bao của một đáp ứng xung rất gồ ghề, đọc RT60 thẳng trên nó thì mỗi người ra một số. Schroeder cho một cách làm nó trơn: tích phân *ngược* bình phương của đáp ứng xung.

$$E(t) = \int_t^{\infty} h^2(\tau)\, d\tau\,; \qquad L(t) = 10 \log_{10} \frac{E(t)}{E(0)} \tag{9.12}$$

Chú ý hệ số 10 chứ không phải 20: E là *năng lượng* chứ không phải biên độ. Nhầm chỗ này thì mọi RT60 đo được đều bằng đúng một nửa giá trị thật — một lỗi im lặng và rất hay gặp.

Từ đường trơn ấy, ISO 3382 định nghĩa cách đo gọi là T30.

$$\mathrm{RT}_{60} = -\frac{60}{s}, \qquad s = \frac{dL}{dt} \text{ với } L \in [-35,\ -5]\ \mathrm{dB} \tag{9.13}$$



(a) Đo RT60 theo cách T30 (b) Đặt bao nhiêu, đo lại bấy nhiêu

Hình 9.25 Đo RT60 theo cách T30, và phép kiểm đặt–đo

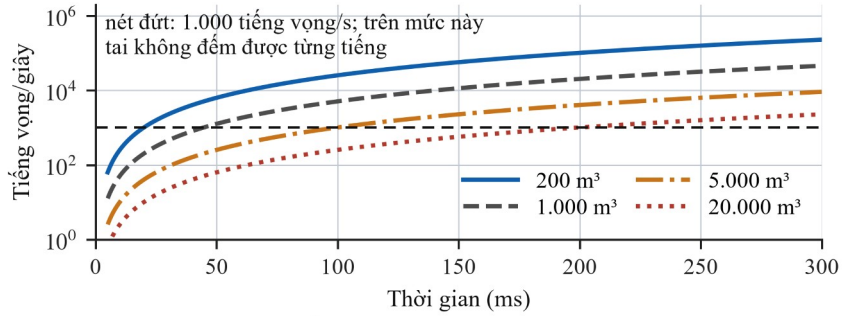
Vì sao không đo thẳng 60 decibel

Vì phần đuôi luôn chìm dưới nhiễu nền của bản thu. Ở −55 decibel thì đường đo được đã là nhiễu chứ không còn là tiếng vang.

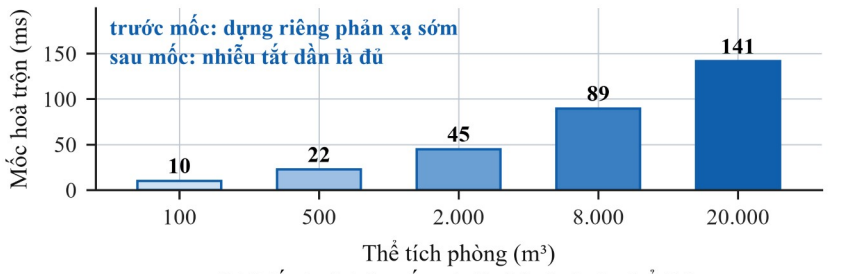
Cách của ISO 3382: lấy đoạn từ −5 tới −35 decibel — đoạn sạch nhất và tuyến tính nhất — khớp một đường thẳng, rồi ngoại suy ra 60 decibel. Bỏ 5 decibel đầu vì ở đó tiếng trực tiếp và phản xạ sớm còn lấn át.

Hình (b) của Hình 9.25 là phép kiểm quan trọng nhất của cả mục: đặt RT60 bao nhiêu thì đo lại đúng bấy nhiêu, lệch nhiều nhất 12 mili giây trên cả sáu điểm. Nghĩa là con số trên giao diện có nghĩa thật chứ không phải một cái núm vô danh.

9.7.5 Mật độ tiếng vọng và mốc hoà trộn



(a) Mật độ tiếng vọng tăng theo bình phương thời gian



(b) Mốc hoà trộn xấp xỉ căn bậc hai của thể tích

Hình 9.26 Mật độ tiếng vọng và mốc hoà trộn theo thể tích phòng

Hai con số này quyết định một bộ vang nghe thật hay nghe giả, và cả hai đều tính được từ vật lý phòng chứ không phải chỉnh mò.

– **Mật độ tiếng vọng tăng theo bình phương thời gian.** Với phòng 1.000 mét khối, tại 50 mili giây đã có chừng 1.300 tiếng vọng mỗi giây; tại 200 mili giây là hơn 20.000. Trên khoảng một nghìn mỗi giây thì tai không đếm được từng tiếng nữa.
– **Mốc hoà trộn xấp xỉ căn bậc hai của thể tích.** Phòng 1.000 mét khối cho 32 mili giây; nhà hát 20.000 mét khối cho 141. Trước mốc ấy phải dựng phản xạ sớm riêng; sau mốc ấy thì nhiễu tắt dần là đủ.

Ba triệu chứng của một bộ vang dựng ẩu

Nghe "lạch cạch" trong vài chục mili giây đầu — mật độ tiếng vọng quá thấp. Chữa bằng cách thêm bộ thông toàn phần.

Nghe có một cao độ ngân lên — bốn độ trễ của bốn bộ lược không nguyên tố cùng nhau, các đỉnh cộng hưởng chồng lên nhau.

Nghe "chói" ở đuôi — thiếu phần nghiêng phổ. Phòng thật bao giờ cũng hút cao tần nhanh hơn trầm, nên đuôi vang phải tối dần đi. Chữa bằng cách đặt một bộ lọc kệ cao âm ngay trong vòng hồi tiếp của mỗi bộ lược.

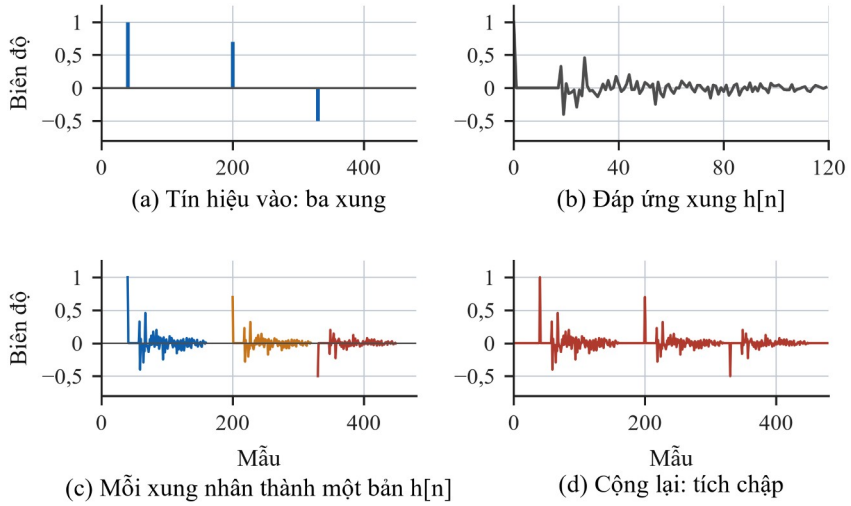
9.8 Tích chập đáp ứng xung

9.8.1 Tích chập là gì — nói bằng hình

Có một cách khác để làm vang, và nó không dựng phòng nào cả: lấy đáp ứng xung *đo thật* trong một phòng có thật, rồi tích chập tín hiệu với nó.

$$y[n] = \sum_{k=0}^{M-1} h[k]\,x[n-k]\,; \qquad Y(f) = H(f)\,X(f) \tag{9.14}$$

Vế trái là định nghĩa trong miền thời gian; vế phải là điều làm cho mọi chuyện khả thi: tích chập trong miền thời gian bằng phép *nhân* trong miền tần số.

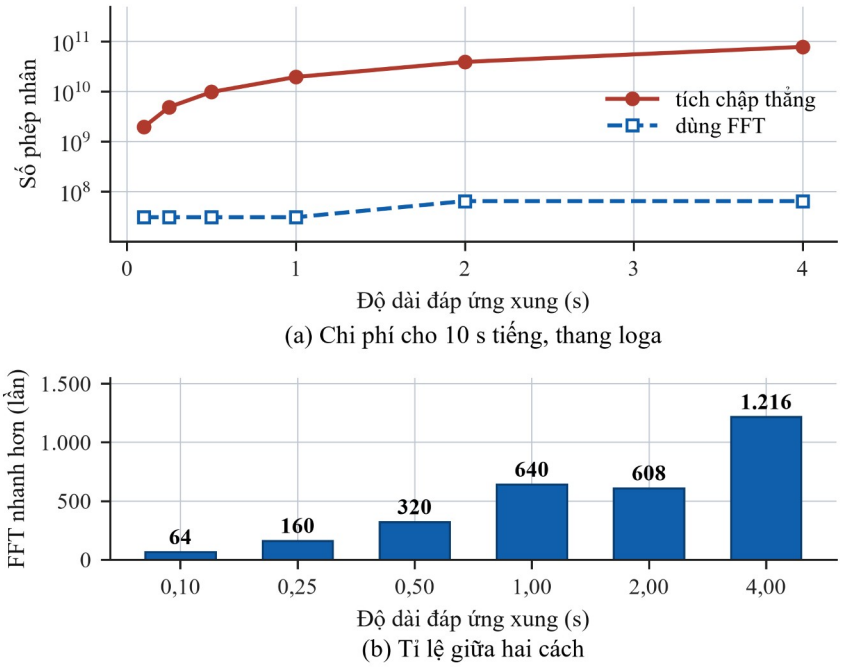


Hình 9.27 Tích chập: thay mỗi mẫu bằng một bản sao của đáp ứng xung

Ý tưởng rất cụ thể: thay *mỗi* mẫu của tín hiệu vào bằng một bản sao của đáp ứng xung, nhân với giá trị mẫu ấy, rồi cộng tất cả lại. Ở hình (a) tín hiệu vào chỉ có ba xung nên đếm được ba bản sao. Với một tệp thật thì có bao nhiêu mẫu là bấy nhiêu bản sao.

9.8.2 Vì sao bắt buộc phải dùng FFT

$$\frac{N \cdot M}{3\,N_{\text{fft}} \log_2 N_{\text{fft}} + N_{\text{fft}}} \tag{9.15}$$



(a) Chi phí cho 10 s tiếng, thang loga

(b) Tỉ lệ giữa hai cách

Hình 9.28 Chi phí tính: tích chập thẳng so với dùng FFT

Với mười giây tiếng và đáp ứng xung hai giây ở 44,1 kilohertz, cách thẳng cần gần bốn mươi tỉ phép nhân; dùng FFT chỉ cần sáu mươi tư triệu — nhanh hơn hơn sáu trăm lần. Trên máy tính hôm nay, một bên mất hàng phút, bên kia mất chưa tới một giây.

Với xử lý thời gian thực thì không phải nhanh hơn — mà là làm được hay không làm được.

Chỗ tích chập gặp lại mục 6.9

Một tệp dài không thể FFT một lần — sẽ hết bộ nhớ. Phải xử lý *theo khối*, đúng như mục 6.9.

Trạng thái mang qua giữa các khối ở đây là *đuôi* của phép tích chập trước: mỗi khối dài L mẫu sinh ra L + M − 1 mẫu, phần thừa phải cộng vào khối sau. Kỹ thuật này gọi là chồng-cộng.

Bỏ sót phần cộng ấy thì mã vẫn chạy và tệp vẫn ghi ra, chỉ là ở mỗi biên khối có một chỗ đứt tiếng vang. Đúng loại lỗi im lặng đã gặp ở chương 6 và chương 8.

Phép kiểm: tích chập theo khối phải cho kết quả *trùng* tích chập một lần, sai khác dưới một phần nghìn tỉ.

Liệt kê 9.4 Tích chập chồng-cộng — mười dòng

```
def tich_chap_theo_khoi(x, h, n_khoi=8192):
    M = h.size
    L = n_khoi
    n_fft = 1 << math.ceil(math.log2(L + M - 1))
    H  = np.fft.rfft(h, n_fft)       # tính MỘT lần
    ra = np.zeros(x.size + M - 1)
```

```
    for i in range(0, x.size, L):
        khoi = x[i:i + L]
        y = np.fft.irfft(np.fft.rfft(khoi, n_fft) * H, n_fft)
        het = min(i + y.size, ra.size)
        ra[i:het] += y[:het - i]     # CỘNG, không phải gán
    return ra
```

9.8.3 Hai loại vang — chọn cái nào

	VANG THUẬT TOÁN	VANG TÍCH CHẬP
Dựng từ	Bộ lọc lược và thông toàn phần	Một đáp ứng xung đo thật trong phòng
Chỉnh được gì	RT60, trễ đầu, mật độ — từng thứ một	Gần như không — chỉ đổi cả tệp
Chi phí tính	Rất rẻ, vài chục phép nhân mỗi mẫu	Đắt; bắt buộc dùng FFT theo khối
Độ trễ	Gần như bằng không	Ít nhất bằng độ dài một khối
Nghe ra	Có thể rất hay, nhưng là một phòng tưởng tượng	Đúng phòng đã đo, tới từng phản xạ sớm
Hợp với	Nhạc — cần chỉnh cho vừa bản dựng	Phim và trò chơi — cần đúng một không gian có thật

Hai cách không thay thế nhau: vang tích chập cho không gian chung, vang thuật toán cho từng lớp, vì mỗi lớp cần một RT60 riêng.

Hình 9.29 Vang thuật toán và vang tích chập — sáu điểm so sánh

Hai cách không thay thế nhau. Trong một bản dựng thật, người ta hay dùng vang tích chập cho không gian chung và vang thuật toán cho từng lớp — vì lớp nào cũng cần một RT60 riêng, mà vang tích chập thì không chỉnh RT60 được.

> **Một lưu ý về chuẩn hoá đáp ứng xung**
>
> Chuẩn hoá đáp ứng xung theo *đỉnh* là sai. Đỉnh của một đáp ứng xung nhà thờ rất nhỏ so với năng lượng tổng của nó, còn đỉnh của một đáp ứng xung phòng nhỏ thì lớn — chuẩn theo đỉnh làm tệp ra to nhỏ thất thường mỗi khi đổi đáp ứng xung.
>
> Cách đúng: chuẩn hoá theo *năng lượng*, tức chia cho căn của tổng bình phương. Khi ấy đổi đáp ứng xung không đổi mức chung.
>
> Hàm chuan_hoa_dap_ung_xung của mô-đun làm đúng việc ấy, và phép kiểm của nó là tổng bình phương phải bằng đúng 1.

9.9 Thực hành, bài tập và tổng kết

9.9.1 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 1 — biquad.** Viết hàm sinh hệ số cho kiểu lọc đỉnh, vẽ đáp ứng ở bốn lượng nâng, rồi kiểm ba điều: tại đúng fc phải cho đúng lượng nâng; ở tần số xa phải cho 0 decibel; cả bốn bộ lọc phải ổn định.

- **Bài 2 — chuỗi EQ.** Dựng một chuỗi bốn dải cho đoạn guitar của người đọc. Cho nhiễu trắng đi qua chuỗi ấy, đo phổ trước và sau, lấy hiệu, rồi so với đường thiết kế. Hai đường phải trùng nhau.
- **Bài 3 — bộ vang.** Dựng bộ vang Schroeder bốn bộ lược và hai bộ thông toàn phần. Đo lại RT60 bằng hàm do_rt60 và kiểm xem có khớp giá trị đặt không. Rồi bỏ hai bộ thông toàn phần đi và *nghe* lại sự khác biệt.

9.9.2 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 9.4 Bảy bài tập của chương 9

#	Đề bài	Đáp số
1	Ở 48 kHz, tính omega không dưới và alpha cho một biquad có fc = 1 kHz và Q = 0,7071. Rồi tính năm hệ số của kiểu thông dải.	omega = 0,1309 rad; alpha = 0,0923
2	Đổi Q = 2 sang bề rộng quãng tám. Rồi đi ngược: muốn bề rộng một phần ba quãng tám thì Q bằng bao nhiêu?	0,714 quãng tám; Q = 4,319
3	Một chuỗi gồm lọc đỉnh +6 dB tại 1 kHz với Q = 1, nối tiếp kệ cao −3 dB tại 8 kHz. Đáp ứng tổng tại 1 kHz và tại 12 kHz là bao nhiêu?	Tại 1 kHz: +6,0 dB. Tại 12 kHz: −2,7 dB
4	Điều chế vòng một tín hiệu 440 Hz với sóng mang 60 Hz. Hai dải bên ở đâu? Vì sao kết quả không còn cao độ 440 Hz?	380 Hz và 500 Hz
5	Một bộ lọc lược của bộ vang có D = 37,1 ms. Muốn RT60 đúng 2,2 giây thì g bằng bao nhiêu? Kiểm lại theo chiều ngược.	g = 0,890; thay ngược ra đúng 2,20 s
6	Tích chập 5 phút tiếng với đáp ứng xung 3 giây ở 48 kHz. So số phép nhân của hai cách và nêu tỉ lệ.	Thẳng: 2,07 nghìn tỉ. FFT: 1.225 triệu. Nhanh hơn 1.693 lần
7	Đo suy giảm của nội suy tuyến tính tại 10 kHz với phần lẻ độ trễ 0; 0,25; 0,5. Giải thích vì sao con số ấy làm flanger nghe "đục".	0,00 · −1,68 · −2,42 dB

9.9.3 Mở rộng cho ai muốn đi xa hơn

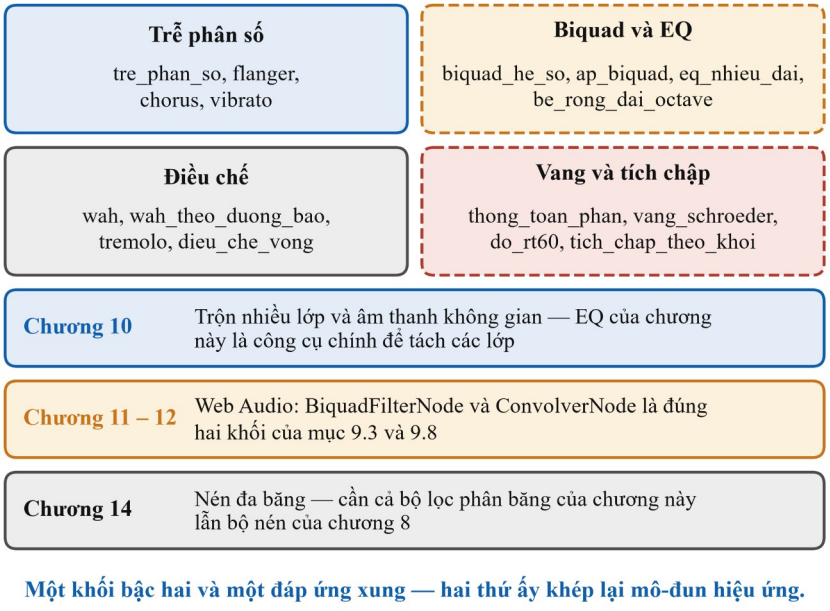
- **EQ động.** Cho lượng nâng của một dải đi theo đường bao mức, đúng như auto-wah của mục 9.5.3 nhưng đổi gain thay vì đổi tần số. Đây là khối chính của một bộ de-esser.
- **Bộ lọc phân băng.** Tách tín hiệu thành ba dải bằng thông thấp, thông dải và thông cao, rồi cộng lại. Kiểm xem tổng có bằng tín hiệu gốc không — câu trả lời là *không*, và lý do nằm ở pha. Đó là bài toán mở đường cho nén đa băng ở chương 14.
- **Nghiêng phổ cho đuôi vang.** Đặt một bộ lọc kệ cao âm ngay *trong* vòng hồi tiếp của mỗi bộ lọc lược, để đuôi vang tối dần đi đúng như phòng thật. So RT60 đo ở dải trầm và dải cao.
- **Đo đáp ứng xung thật.** Phát một tiếng quét hình sin từ thấp lên cao trong một phòng, thu lại, rồi giải chập để lấy đáp ứng xung. Đây là cách các thư viện đáp ứng xung thương mại được làm ra, và làm được với thiết bị rất đơn giản.

9.9.4 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 nội suy tuyến tính làm flanger nghe "đục", trong khi cùng phép nội suy ấy dùng cho một độ trễ *cố định* thì không sao?
- Vì sao bậc hai là bậc thấp nhất dựng được một bộ lọc có cộng hưởng? Bậc một thiếu cái gì?
- Với bộ lọc kệ, tại đúng fc chỉ đi được nửa lượng nâng. Với lọc đỉnh thì tại fc cho đúng lượng nâng. Vì sao khác nhau?
- Vì sao thứ tự các dải trong một chuỗi EQ không quan trọng, trong khi thứ tự nén rồi giới hạn ở chương 8 thì rất quan trọng?
- Tremolo và điều chế vòng cùng là một phép nhân. Chỉ ra hai khác biệt, và nói rõ khác biệt nào làm sóng mang biến mất.
- Vì sao Schroeder dùng bộ thông toàn phần chứ không dùng thêm bộ lọc lược để làm dày mật độ tiếng vọng?
- Trong công thức đường tắt năng lượng, vì sao là $10 \cdot \log_{10}$ chứ không phải $20 \cdot \log_{10}$? Nhầm thì RT60 đo được sai bao nhiêu?

9.9.5 Tổng kết — chương 9 để lại gì



Hình 9.30 Chương 9 để lại gì cho các chương sau

Bảng 9.5 Mười ba hàm mới của chương 9

Hàm	Làm gì	Dùng lại ở
tre_phan_so	Trễ với độ trễ là số thực, ba kiểu nội suy	Chương 10 — panning theo trễ
flanger, chorus, vibrato	Ba hiệu ứng điều biến, khác nhau ở đường độ trễ	Chương 11 – 12 — rack hiệu ứng
biquad_he_so, ap_biquad	Sinh năm hệ số từ ba con số, và chạy bộ lọc	Chương 10, 11, 14 — dùng ở khắp nơi
dap_ung_biquad_db, cuc_biquad, on_dinh_biquad	Đọc và kiểm một bộ lọc trước khi chạy nó	Mọi chương còn lại
eq_nhieu_dai, dap_ung_eq_db	Chuỗi EQ và đáp ứng tổng của nó	Chương 10 — tách lớp trong bản trộn
be_rong_dai_octave, q_tu_be_rong	Đổi qua lại giữa Q và quãng tám	Chương 14
wah, wah_theo_duong_bao	Bộ thông dải có tần số chạy	Chương 14 — de-esser và EQ động
tremolo, dieu_che_vong	Hai phép điều chế biên độ	Chương 11 – 12
thong_toan_phan	Biên độ phẳng, chỉ đổi pha	Chương 10 — mạng trễ phản hồi
vang_schroeder	Bộ vang bốn lược, hai thông toàn phần	Chương 10 và chương 14
duong_nang_luong_db, do_rt60	Đo RT60 theo cách T30 của ISO 3382	Chương 10
tich_chap_nhanh, tich_chap_theo_khoi	Tích chập bằng FFT, một lần và theo khối	Chương 11 – 12 — ConvolverNode
dap_ung_xung_gia_lap, chuan_hoa_dap_ung_xung	Dựng và chuẩn hoá đáp ứng xung	Chương 10

Ba câu mang về

Cả chương chỉ có *một* khối mới: bộ lọc bậc hai. Năm mục còn lại là cách dùng nó, hoặc cách ghép nó với khối trễ của chương 8.

Ba con số — tần số, lượng nâng, Q — trả lời ba câu độc lập nhau. Chỉnh EQ mà không trả lời được ba câu ấy thì chỉ là đoán, giống như chỉnh núm bộ nén ở chương 8.

Ba chỗ có phép kiểm bằng số, và cả ba đều nên chạy: nội suy trễ mất 2,42 decibel ở phần lẻ 0,5; chuỗi EQ đo lại trùng thiết kế dưới 0,01 decibel; bộ vang đặt RT60 bao nhiêu thì đo lại đúng bấy nhiêu, lệch dưới 12 mili giây.

Chương 10 mở sang phần trộn và không gian: headroom khi chồng nhiều lớp, panning, hiệu ứng Haas, rồi binaural và HRTF. Bộ EQ của chương này là công cụ chính để tách các lớp ra khỏi nhau, còn bộ vang thì trở thành cách đặt các lớp

vào cùng một không gian. Cần chuẩn bị trước: bốn tới sáu tệp WAV của cùng một bản nhạc tách theo nhạc cụ, và một đôi tai nghe kín.

Tài liệu tham khảo của chương

Nguồn trực tiếp của các công thức trong bài

- Bristow-Johnson, R. — Cookbook formulae for audio EQ biquad filter coefficients (thường gọi tắt là "RBJ Audio EQ Cookbook"). Đây là nguồn *trực tiếp* của toàn bộ mục 9.3 và 9.4: bảy kiểu lọc, công thức của omega không dưới và alpha, và cách chuẩn hoá hệ số. Tài liệu chỉ vài trang và là thứ đáng in ra dán lên tường.
- Schroeder, M. R. (1962) — Natural Sounding Artificial Reverberation, Journal of the Audio Engineering Society 10(3). Bài báo gốc của sơ đồ ở mục 9.7.2, gồm cả lập luận vì sao phải dùng bộ thông toàn phần.
- Schroeder, M. R. (1965) — New Method of Measuring Reverberation Time. Nguồn của đường tích phân ngược ở mục 9.7.4.
- ISO 3382-1 — Acoustics: Measurement of room acoustic parameters. Nguồn của cách đo T30 và của khoảng −5 tới −35 decibel.
- Laakso, T. và cộng sự (1996) — Splitting the Unit Delay: Tools for Fractional Delay Filter Design, IEEE Signal Processing Magazine. Nguồn của mục 9.2, gồm cả công thức đáp ứng của nội suy tuyến tính và họ nội suy Lagrange.

Sách

- Zölzer, U. (chủ biên) — DAFX: Digital Audio Effects. Chương về bộ lọc, về hiệu ứng điều biến và về vang tương ứng đúng các mục 9.2 tới 9.8 của chương này.
- Smith, J. O. — Introduction to Digital Filters, và Physical Audio Signal Processing. Nguồn chặt chẽ nhất cho phần cực–không và cho bộ thông toàn phần. Đọc miễn phí trên trang của tác giả tại Đại học Stanford.
- Välimäki, V. và cộng sự (2012) — Fifty Years of Artificial Reverberation, IEEE Transactions on Audio, Speech and Language Processing. Bài tổng quan đọc được sau khi đã nắm mục 9.7.
- Reiss, J. D., McPherson, A. — Audio Effects: Theory, Implementation and Application. Viết cho người lập trình, nhiều mã giả.

Công cụ nên biết

- scipy.signal — có sẵn hàm butter, cheby1, sosfilt và freqz. Dùng để đối chiếu kết quả của mô-đun tự viết. Lưu ý scipy dùng dạng "sos" (mảng các khối bậc hai) chứ không dùng b, a rời — chính vì lý do ổn định số học nêu ở mục 9.3.2.
- Một bộ EQ có hiển thị đường đáp ứng — bản miễn phí đi kèm Reaper hoặc Ardour đều đủ. Điều quan trọng là *nhìn* được đường cong khi chỉnh từng núm.

- Thư viện đáp ứng xung miễn phí: OpenAIR của Đại học York, và bộ EchoThief. Dùng cho phần tích chập ở mục 9.8.
- ffmpeg với bộ lọc afir và aecho — để đối chiếu vang tích chập và vang thuật toán từ dòng lệnh.

Đọc thêm theo hướng

- **Muốn đi sâu về bộ lọc:** tìm đọc về phép biến đổi song tuyến và về "warping" tần số — đó là lý do một bộ lọc số thiết kế theo công thức tương tự bị lệch tần số ở gần Nyquist, và cách bù lại.
- **Muốn đi sâu về vang:** tìm đọc về mạng trễ phản hồi (feedback delay network) của Jot và Chaigne. Đó là bước tiến lớn sau Schroeder, và nó dùng một ma trận trực giao để trộn các đường trễ với nhau.
- **Muốn đi sâu về tích chập:** tìm đọc về tích chập phân khối (partitioned convolution) của Gardner. Nó chia đáp ứng xung thành nhiều đoạn dài dần, để có độ trễ thấp ở đầu mà vẫn rẻ ở đuôi — đúng cách các bộ vang tích chập chạy thời gian thực làm.

Một lưu ý về các con số trong bài

> **Cách đọc các con số**
>
> Các con số về đáp ứng bộ lọc, về quy đổi Q sang quãng tám, về vị trí dải bên của điều chế vòng và về chi phí tích chập là kết quả tính chính xác từ công thức — chúng đúng với mọi tư liệu.
>
> Các con số về tư liệu (hợp âm guitar, nhịp trống) là kết quả *đo* trên tư liệu dựng bằng mã, mô tả trong chú thích từng hình.
>
> Ba phép kiểm đặt–đo trong bài — nội suy trễ, chuỗi EQ, và RT60 — nên tự chạy lại. Chúng là cách duy nhất để biết mã do người đọc viết có đúng hay không, vì cả ba lỗi tương ứng đều *im lặng*.
>
> Mọi phép tính trong bài sinh ra từ mô-đun sigtools.py, đã qua 329 phép tự kiểm.