# MỤC LỤC

**Chương 1. Nhập môn và bản chất vật lý của âm thanh**.....5
1.1 Vị trí của chương.....5
1.2 Âm thanh là sóng cơ học.....7
1.3 Bốn đại lượng gốc.....11
1.4 Phổ hài và âm sắc.....17
1.5 Tai người nghe như thế nào.....23
1.6 Hướng dẫn thực hành.....30
1.7 Bài tập tính toán.....32
1.8 Câu hỏi ôn tập và bài thảo luận.....34
1.9 Tóm tắt và chuẩn bị cho chương sau.....37

# DANH MỤC HÌNH VẼ

Hình 1.1 Bản đồ mười lăm chương của cuốn sách.....6
Hình 1.2 Bốn câu hỏi của chương 1 và chỗ chúng quay lại trong cuốn sách.....6
Hình 1.3 Ba yếu tố của âm thanh: nguồn, môi trường và nơi nhận.....7
Hình 1.4 Có môi trường thì âm truyền được; trong chân không thì không.....7
Hình 1.5 Sóng dọc trong không khí — vùng nén và vùng giãn, cùng đường cong áp suất tương ứng.....8
Hình 1.6 Áp suất tại một điểm theo thời gian — đây chính là thứ mà micro ghi lại.....9
Hình 1.7 Tốc độ âm theo nhiệt độ không khí, và trong vài môi trường khác.....10
Hình 1.8 Bước sóng trên toàn dải nghe được — trải hơn một nghìn lần.....11
Hình 1.9 Bốn đại lượng gốc, đọc hết được trên một dạng sóng.....12
Hình 1.10 Ba tần số cách nhau đúng một quãng tám, nhìn ở miền thời gian và miền tần số.....12
Hình 1.11 Ba tín hiệu cùng đỉnh 1,0 nhưng năng lượng và hệ số đỉnh khác hẳn nhau. .14
Hình 1.12 Hai thang decibel: hệ số 20 cho biên độ, hệ số 10 cho công suất.....15
Hình 1.13 Mỗi lần gấp đôi khoảng cách thì mức giảm đúng 6 dB.....16
Hình 1.14 Hai sóng giống hệt nhau, chỉ khác pha: tổng có thể gấp đôi hoặc biến mất..16
Hình 1.15 Ba dạng sóng cùng tần số cơ bản 300 Hz và ba phổ tương ứng.....17
Hình 1.16 Cộng dần các hài lẻ: từ một sóng sin tới sóng vuông.....18
Hình 1.17 Ba âm sắc khác nhau dựng từ cùng một tần số cơ bản, chỉ khác tỉ lệ các hài .....18
Hình 1.18 Đường bao biên độ ADSR của một nốt: bốn giai đoạn và bốn tham số.....19
Hình 1.19 Cùng một dạng sóng, ba đường bao khác nhau: gõ hoặc gảy, kéo vĩ, thổi (tần số sóng được hạ thấp để nhìn rõ).....21
Hình 1.20 Phổ đồ của một tín hiệu quét tần số: dạng sóng không cho biết gì, phổ đồ thì nói hết.....23
Hình 1.21 Dải nghe được của tai người và chỗ đứng của các nguồn âm quen thuộc.....24
Hình 1.22 Ngưỡng nghe tuyệt đối — mức cần có để nghe thấy, theo từng tần số.....24
Hình 1.23 Đường đẳng âm theo ISO 226:2003: mỗi đường nối các âm nghe to bằng nhau, đánh số bằng phon.....26
Hình 1.24 Quan hệ giữa phon và sone: từ 40 phon trở lên, thêm 10 phon thì độ to gấp đôi (thang sone là thang loga).....27
Hình 1.25 Che lấp theo tần số: một âm mạnh nâng ngưỡng nghe của vùng lân cận.....29
Hình 1.26 Từ một hiện tượng sinh lý tới định dạng nén có tổn hao.....29
Hình 1.27 Môi trường làm việc dùng chung cho cả mười lăm chương.....30
Hình 1.28 Dạng sóng và phổ phải thấy sau khi chạy đoạn mã kiểm môi trường.....32
Hình 1.29 Bản đồ khái niệm chương 1 — bốn nhánh quay lại ở mọi chương sau.....37

# DANH MỤC BẢNG BIỂU

Bảng 1.1 Tốc độ âm trong không khí ở vài mức nhiệt độ....9
Bảng 1.2 Bước sóng ở vài tần số tiêu biểu, tính ở 20 °C....10
Bảng 1.3 Nốt La ở các quãng tám và tần số tương ứng....13
Bảng 1.4 Vài mốc decibel cần thuộc....15
Bảng 1.5 Đặc điểm phổ hài của vài nhóm nhạc cụ....19
Bảng 1.6 Bốn giai đoạn của đường bao ADSR....20
Bảng 1.7 Đường bao điển hình của vài nhóm nhạc cụ (mức độ gần đúng, mang tính minh hoạ)....21
Bảng 1.8 Mức áp suất âm (dB SPL) cần có để nghe to bằng nhau, theo ISO 226:2003 26
Bảng 1.9 Vài mốc phon và sone (ví dụ chỉ mang tính minh hoạ)....28
Bảng 1.10 Triệu chứng, nguyên nhân và cách sửa....32
Bảng 1.11 Thuật ngữ chương 1....37

# CHƯƠNG 1
# NHẬP MÔN VÀ BẢN CHẤT VẬT LÝ CỦA ÂM THANH

**Mục tiêu của chương**

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

- Giải thích được âm thanh là gì, cần những yếu tố nào để tồn tại, và vì sao nó không truyền được trong chân không.
- Gọi tên và giải thích được bốn đại lượng gốc của sóng âm, cùng quan hệ giữa chúng với cảm nhận của tai.
- Tính được bước sóng, tốc độ âm theo nhiệt độ, mức decibel và mức suy giảm theo khoảng cách.
- Giải thích được vì sao cùng một nốt mà mỗi nhạc cụ nghe khác nhau, bằng khái niệm phổ hài.
- Mô tả được hai điểm mù của tai người và nêu được vì sao chúng là nền của nén có tổn hao.
- Cài xong môi trường lập trình và chạy được đoạn mã sinh, đo, ghi và vẽ một sóng sin.

> **Một điều về cách viết sách này**
>
> Mọi *hình* trong tài liệu đều được dựng bằng mã từ tín hiệu tính thật, không phải ảnh sưu tầm. Dạng sóng, phổ, phổ đồ, đường cong ngưỡng nghe — tất cả đều tính ra rồi vẽ.
>
> Mọi *con số* trong bài và mọi *đáp số* của bài tập đều do một mô-đun tính toán sinh ra, và mô-đun ấy có bộ tự kiểm chạy được. Đề bài đổi tham số thì đáp số tự đổi theo, không bao giờ lệch nhau.
>
> Chính mô-đun ấy sẽ được trao cho người đọc từ chương 5, khi bắt đầu tự viết thư viện của riêng mình.

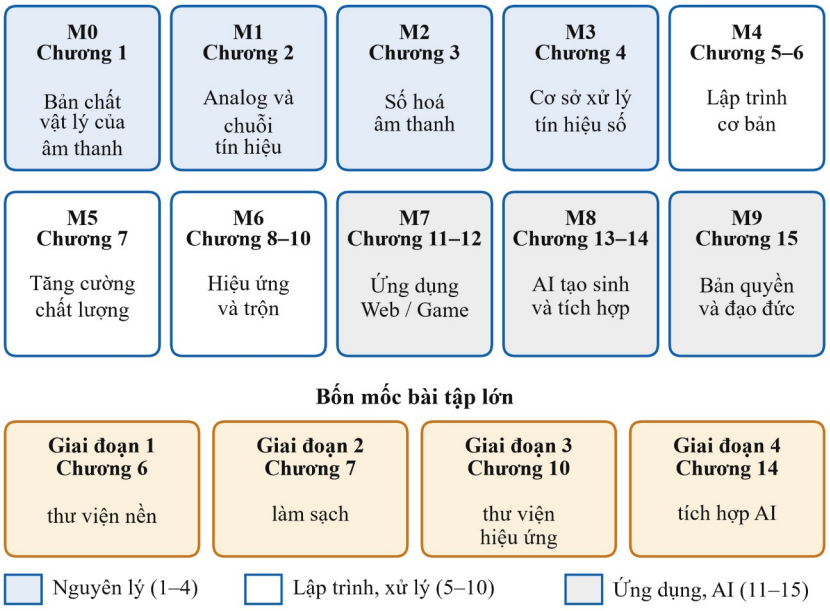
## 1.1 Vị trí của chương

### 1.1.1 Cuốn sách này dạy gì

Lập trình âm thanh là cuốn sách dạy cách viết mã xử lý tín hiệu âm thanh. Đọc xong cuốn sách, người đọc tự viết được một thư viện xử lý âm thanh của riêng mình, và đưa được thư viện ấy vào một sản phẩm chạy được — một ứng dụng web hoặc một trò chơi.

Chữ quan trọng nhất trong câu trên là *tự viết*. Ngày nay có sẵn rất nhiều thư viện làm hộ mọi việc, và có cả công cụ sinh mã giúp. Nhưng người chỉ biết gọi thư viện thì dừng lại ngay khi thư viện không có sẵn hàm cần dùng, hoặc khi kết quả

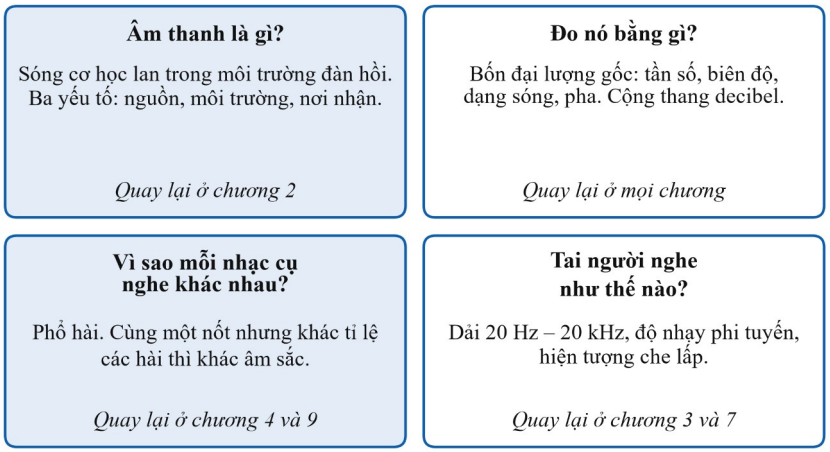
sai mà không biết sai ở đâu. Cuốn sách này đi theo hướng ngược lại: viết lấy phần lõi, hiểu từng dòng, rồi mới dùng thư viện cho phần ngoại vi.




*Hình 1.1 Bản đồ mười lăm chương của cuốn sách*

Nhìn bản đồ trên sẽ thấy bốn chương đầu — trong đó có chương này — chưa viết dòng mã nào. Đó là chủ ý. Bốn chương ấy dựng phần *nguyên lý*, và mọi dòng mã viết từ chương 5 trở đi đều là cách diễn đạt lại một nguyên lý trong bốn chương ấy bằng ngôn ngữ máy tính. Ai bỏ qua phần này thì về sau chỉ chép được mã chứ không sửa được mã.

### 1.1.2 Bốn câu hỏi của chương này




*Hình 1.2 Bốn câu hỏi của chương 1 và chỗ chúng quay lại trong cuốn sách*

Chương được tổ chức quanh bốn câu hỏi ở hình trên. Dòng cuối mỗi ô cho biết kiến thức ấy sẽ quay lại ở chương nào — đây không phải chi tiết trang trí mà là điều đáng để ý: không có phần nào của chương này chỉ học một lần rồi bỏ.

**Cách dùng chương này**

Không cần thuộc lòng con số nào. Cần hiểu *quan hệ* giữa các đại lượng: đổi cái này thì cái kia đổi theo ra sao.

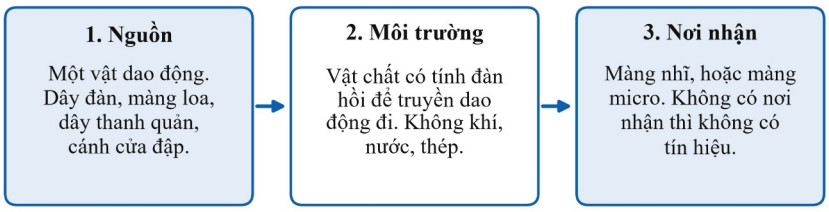
Mọi hình trong chương này đều được vẽ từ tín hiệu tính thật bằng mã. Cuối chương sẽ có đoạn mã để tự dựng lại vài hình trong số đó.

Phần thính giác ở cuối chương là phần bị coi nhẹ nhất và cũng là phần có ích nhất — nó là nền của toàn bộ chuyện nén ở chương 3.

## 1.2 Âm thanh là sóng cơ học

### 1.2.1 Ba yếu tố phải có đủ

Âm thanh không phải một vật. Nó là một *hiện tượng*, và hiện tượng ấy chỉ xảy ra khi có đủ ba yếu tố.

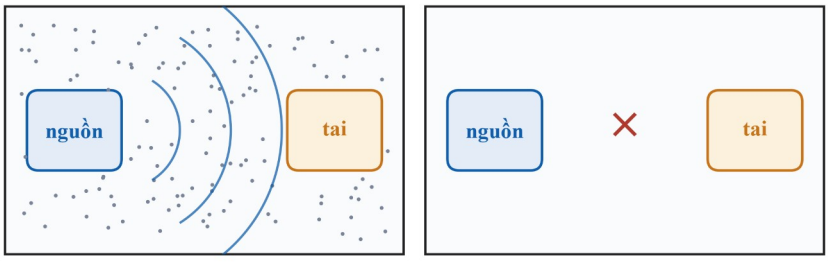



*Thiếu môi trường thì không có âm thanh — đó là lý do trong chân không hoàn toàn im lặng.*

*Hình 1.3 Ba yếu tố của âm thanh: nguồn, môi trường và nơi nhận*

- **Nguồn** — một vật đang dao động. Dây đàn được gảy, màng loa bị dòng điện đẩy tới đẩy lui, dây thanh quản rung khi chúng ta nói.
- **Môi trường** — vật chất có tính đàn hồi để dao động ấy truyền đi được. Không khí là môi trường quen thuộc nhất, nhưng nước, gỗ và thép đều truyền âm, và truyền tốt hơn không khí nhiều.
- **Nơi nhận** — màng nhĩ, hoặc màng của một chiếc micro. Không có nơi nhận thì dao động vẫn lan đi, nhưng không có tín hiệu nào được sinh ra.

Yếu tố thứ hai là yếu tố dễ bị bỏ sót, và nó dẫn tới một hệ quả dễ kiểm chứng.




(a) Có không khí — âm truyền được

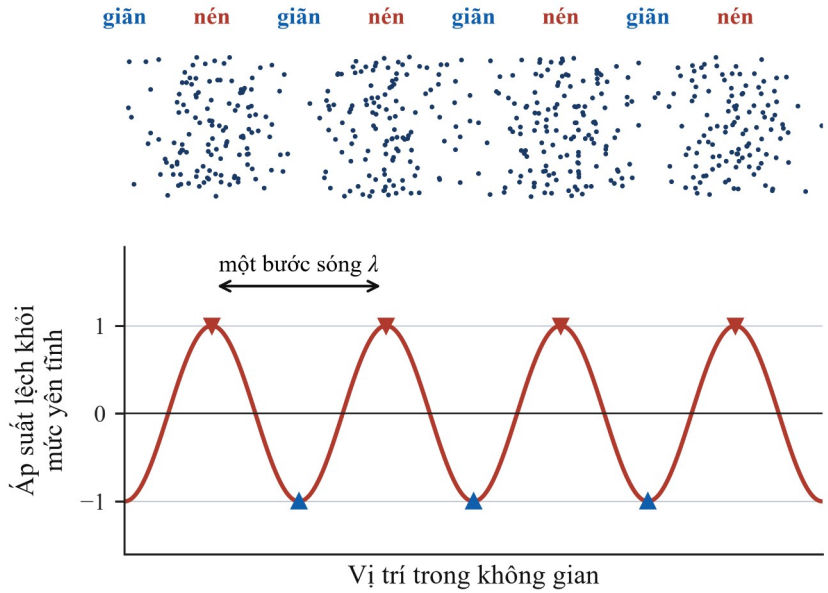
(b) Chân không — không có gì để truyền

*Hình 1.4 Có môi trường thì âm truyền được; trong chân không thì không*

Đây là chỗ phân biệt âm thanh với ánh sáng. Ánh sáng là sóng điện từ nên đi được trong chân không — vì thế chúng ta nhìn thấy Mặt Trời. Âm thanh là sóng *cơ học* nên bắt buộc phải có vật chất để truyền. Mọi cảnh phim vũ trụ có tiếng nổ đều là quy ước nghệ thuật chứ không phải vật lý.

### 1.2.2 Sóng dọc: cái lan đi là vùng nén, không phải các phần tử

Khi một vật dao động trong không khí, nó đẩy các phần tử khí ngay cạnh nó. Các phần tử ấy dồn lại, rồi đẩy tiếp lớp kế bên, rồi bật trở về chỗ cũ. Kết quả là một vùng *nén* và một vùng *giãn* nối tiếp nhau chạy đi trong không gian.




*Hình 1.5 Sóng dọc trong không khí — vùng nén và vùng giãn, cùng đường cong áp suất tương ứng*

**Điều dễ hiểu sai nhất về sóng âm**

Các phần tử không khí *không* đi theo sóng. Chúng chỉ rung tới lui quanh vị trí cũ của chúng.

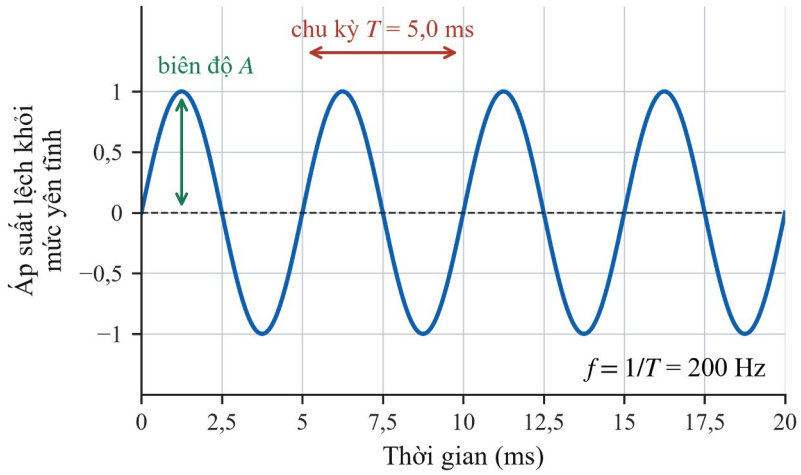
Cái lan đi trong không gian là *trạng thái nén*, không phải vật chất. Giống như sóng chạy trên một hàng người xếp hàng: sóng đi từ đầu hàng tới cuối hàng, nhưng không ai đổi chỗ.

Nếu các phần tử đi theo sóng thì mỗi lần chúng ta nói sẽ có một luồng gió thổi tới người nghe — điều đó không xảy ra.

Sóng âm trong không khí là sóng *dọc*: phần tử dao động theo đúng hướng sóng đi. Khác với sóng trên mặt nước, vốn là sóng ngang — phần tử nước nhấp nhô lên xuống trong khi sóng chạy ngang. Phân biệt này quan trọng khi học về phản xạ và giao thoa ở các chương sau.

### 1.2.3 Từ không gian sang thời gian

Hình ở mục trước nhìn sóng theo *không gian* — chụp lại một khoảnh khắc và xem áp suất phân bố thế nào dọc đường đi. Nhưng micro thì không làm thế. Micro đứng yên ở *một* điểm và ghi lại áp suất tại điểm ấy thay đổi ra sao theo *thời gian*.




*Hình 1.6 Áp suất tại một điểm theo thời gian — đây chính là thứ mà micro ghi lại*

Đường cong này là hình quan trọng nhất của cả chương, và có lẽ của cả cuốn sách. Mọi tệp âm thanh trên máy tính đều là đường cong ấy, ghi lại thành một dãy số. Toàn bộ những gì cuốn sách này làm — lọc, nén, thêm hiệu ứng, trộn nhiều lớp — đều chỉ là biến đổi dãy số đó.

Dạng đơn giản nhất của đường cong ấy là hình sin, và nó viết được bằng một công thức bốn tham số:

$$x(t) = A\sin(2\pi ft + \varphi) \tag{1.1}$$

trong đó A là biên độ, f là tần số tính bằng héc, t là thời gian tính bằng giây, và φ là pha ban đầu. Bốn tham số này là toàn bộ thông tin của một sóng sin — mục 1.3 sẽ đi vào từng cái.

### 1.2.4 Tốc độ truyền âm

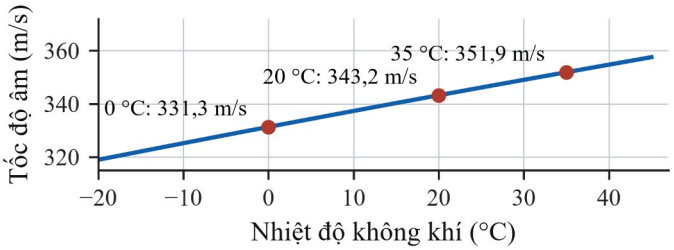
Sóng âm cần thời gian để đi từ nguồn tới tai. Tốc độ ấy phụ thuộc vào *môi trường*, và với không khí thì phụ thuộc thêm vào *nhiệt độ*.

$$c \approx 331{,}3\sqrt{1 + \frac{t}{273{,}15}} \quad \text{[m/s]} \tag{1.2}$$

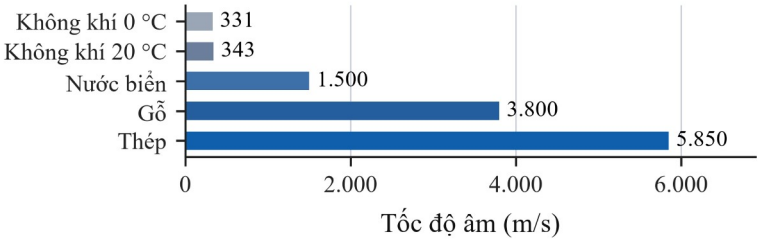
với t là nhiệt độ không khí tính bằng độ C. Công thức này cho những con số quen thuộc sau:

*Bảng 1.1 Tốc độ âm trong không khí ở vài mức nhiệt độ*

| Nhiệt độ | Tốc độ âm | Ghi chú |
|---|---|---|
| 0 °C | 331,3 m/s | giá trị hay gặp trong sách giáo khoa |
| 20 °C | 343,2 m/s | nhiệt độ phòng — con số dùng trong cả cuốn sách này |
| 35 °C | 351,9 m/s | ngoài trời mùa hè ở Hà Nội |




(a) Không khí: phụ thuộc nhiệt độ




(b) Môi trường càng đặc, càng cứng thì âm đi càng nhanh

*Hình 1.7 Tốc độ âm theo nhiệt độ không khí, và trong vài môi trường khác*

Chênh lệch giữa 0 °C và 35 °C là khoảng 20,6 m/s, tức chừng 6,2 phần trăm. Với công việc trong nhà thì lấy tròn 343 m/s là đủ. Nhưng khi căn giờ trễ cho một hệ thống loa ngoài trời, chênh lệch ấy đủ để làm lệch vài mili giây trên quãng đường vài chục mét — và vài mili giây là đủ để nghe thấy.

**Vì sao âm đi nhanh hơn trong chất rắn**

Trực giác thường nói ngược: chất rắn đặc hơn, chắc âm phải đi chậm hơn. Thực tế ngược lại — âm đi trong thép nhanh gấp mười bảy lần trong không khí.

Lý do là *độ cứng*. Tốc độ âm tăng theo độ cứng và giảm theo khối lượng riêng. Thép nặng hơn không khí vài nghìn lần, nhưng cứng hơn nhiều triệu lần — nên phần cứng thắng.

Hệ quả thực hành: âm truyền qua tường và sàn nhà rất tốt. Đó là lý do cách âm một phòng thu khó hơn nhiều so với tưởng tượng — chặn đường không khí là chưa đủ.

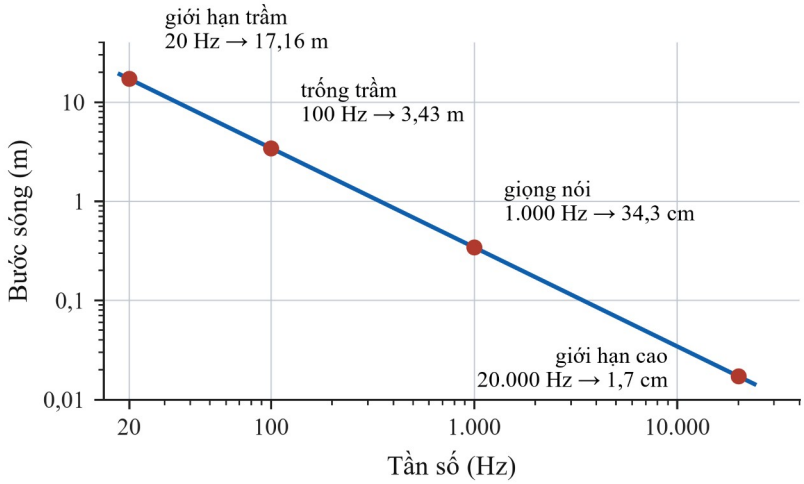
### 1.2.5 Bước sóng — con số quyết định kích thước mọi thứ

Biết tốc độ và tần số thì tính được *bước sóng*: quãng đường sóng đi được trong đúng một chu kỳ.

$$\lambda = \frac{c}{f}; \qquad f = \frac{c}{\lambda} \tag{1.3}$$

*Bảng 1.2 Bước sóng ở vài tần số tiêu biểu, tính ở 20 °C*

| Tần số | Bước sóng | So sánh cho dễ hình dung |
|---|---|---|
| 20 Hz | 17,16 m | dài hơn một phòng học |
| 100 Hz | 3,43 m | chiều cao một tầng nhà |
| 1.000 Hz | 34,3 cm | chiều dài một cẳng tay |
| 10.000 Hz | 3,4 cm | chiều rộng một ngón tay |
| 20.000 Hz | 1,7 cm | bề dày một cục tẩy |



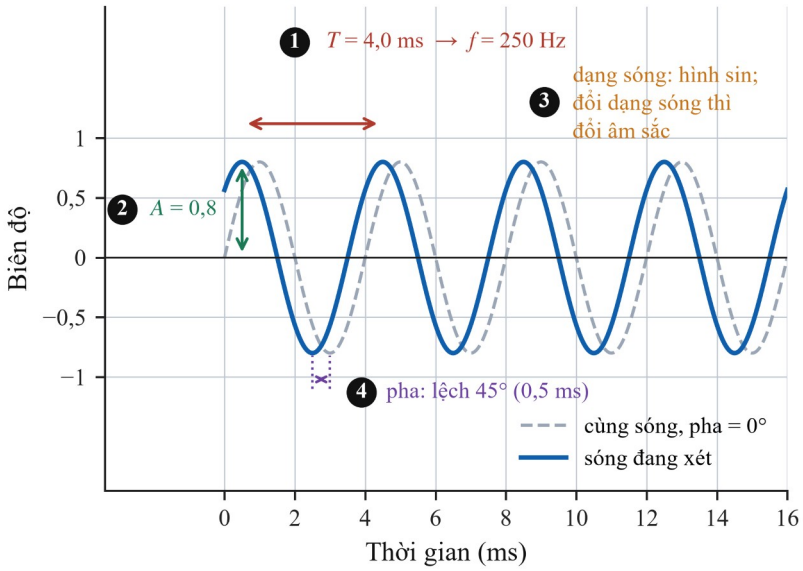

*Hình 1.8 Bước sóng trên toàn dải nghe được — trải hơn một nghìn lần*

Bảng và hình trên giải thích rất nhiều chuyện tưởng như không liên quan. Vì sao loa trầm phải to còn loa cao thì bé? Vì màng loa cần kích thước tương xứng với bước sóng nó phát ra. Vì sao tấm mút tiêu âm mỏng dán tường chỉ hút được âm cao mà không đụng gì tới âm trầm? Vì âm trầm có bước sóng hàng mét, tấm mút dày vài xăng-ti-mét không phải là vật cản đối với nó.

> **Quy tắc bỏ túi**
>
> Một vật chỉ ảnh hưởng đáng kể tới sóng âm khi kích thước của nó *cùng bậc* với bước sóng.
>
> Nhỏ hơn bước sóng nhiều thì sóng đi vòng qua như không có gì.
>
> Đó là lý do đứng sau một cái cột vẫn nghe rõ tiếng nói, nhưng nấp sau cùng cái cột ấy thì không nhìn thấy người — bước sóng ánh sáng quá nhỏ nên không đi vòng được.

## 1.3 Bốn đại lượng gốc

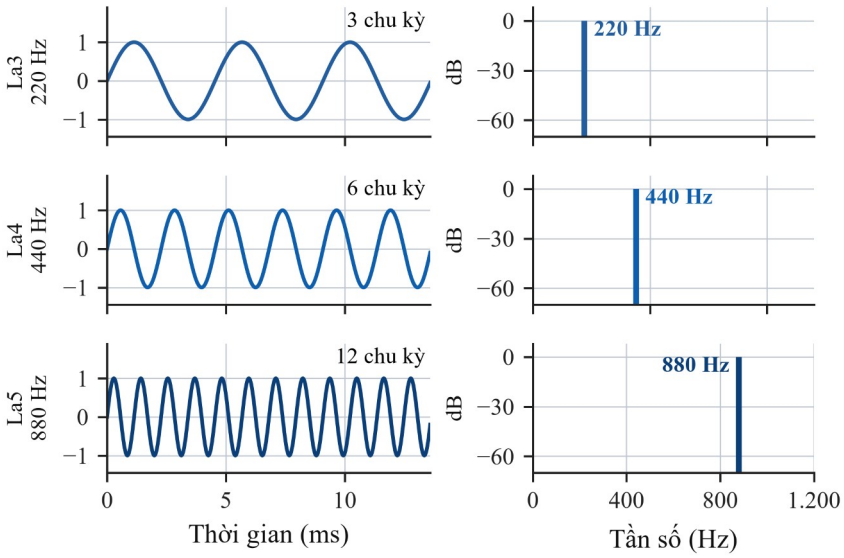
Một sóng sin được xác định hoàn toàn bởi bốn con số. Bốn con số ấy là thứ mà mọi phép xử lý trong cuốn sách này tác động vào, nên phải nắm thật chắc từ chương đầu.




*Hình 1.9 Bốn đại lượng gốc, đọc hết được trên một dạng sóng*

### 1.3.1 Tần số — quyết định cao độ

Tần số là số chu kỳ trọn vẹn trong một giây, đơn vị héc (Hz). Nó là nghịch đảo của chu kỳ: f = 1/T. Tai người quy tần số thành cảm giác *cao độ* — tần số càng lớn thì nghe càng cao.




(a) Dạng sóng, cùng khoảng thời gian: càng cao càng nhiều chu kỳ (b) Phổ: một vạch, đúng chỗ

*Hình 1.10 Ba tần số cách nhau đúng một quãng tám, nhìn ở miền thời gian và miền tần số*

Quan hệ giữa tần số và cao độ là quan hệ *nhân* chứ không phải quan hệ *cộng*. Gấp đôi tần số thì lên đúng một quãng tám, dù đang ở chỗ nào trên thang. Từ 100 lên 200 Hz là một quãng tám; từ 1.000 lên 2.000 Hz cũng là một quãng tám, dù chênh lệch tuyệt đối lớn gấp mười lần.

*Bảng 1.3 Nốt La ở các quãng tám và tần số tương ứng*

| Nốt | Tần số | Quan hệ |
|---|---|---|
| La2 | 110 Hz | thấp hơn hai quãng tám |
| La3 | 220 Hz | thấp hơn một quãng tám |
| La4 | 440 Hz | nốt chuẩn để lên dây đàn |
| La5 | 880 Hz | cao hơn một quãng tám |
| La6 | 1.760 Hz | cao hơn hai quãng tám |

**Vì sao thang tần số hay vẽ theo lô-ga-rít**

Vì tai nghe theo tỉ lệ chứ không theo hiệu. Khoảng cách cảm nhận giữa 100 và 200 Hz bằng đúng khoảng cách giữa 5.000 và 10.000 Hz.

Vẽ trên trục thẳng thì cả dải trầm bị dồn vào một góc bé xíu, còn trên trục lô-ga-rít thì mỗi quãng tám chiếm cùng bề rộng.

Từ nay về sau, mọi biểu đồ phổ trong sách đều dùng trục lô-ga-rít khi cần nhìn cả dải, và trục thẳng khi cần nhìn kỹ một vùng hẹp.

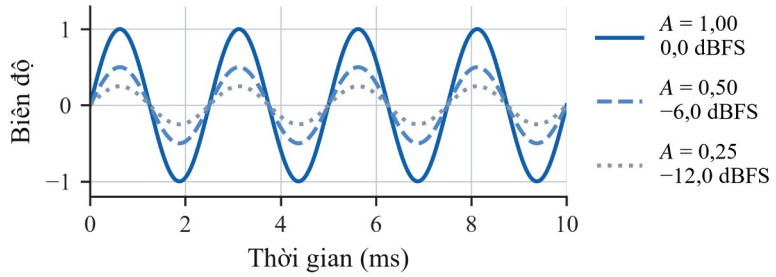
### 1.3.2 Biên độ — quyết định độ to

Biên độ là độ lệch lớn nhất của áp suất khỏi mức yên tĩnh. Trong máy tính, mẫu âm thanh thường được biểu diễn bằng số thực trong khoảng từ −1,0 tới +1,0; biên độ 1,0 gọi là *toàn thang*.

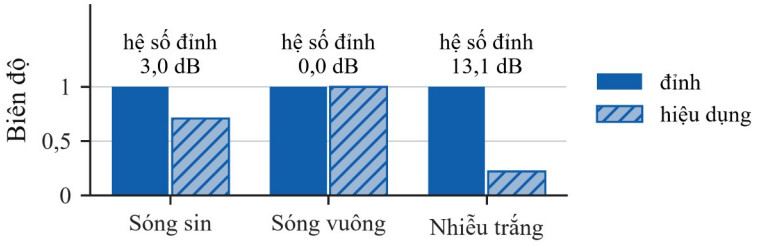
Nhưng biên độ đỉnh không phải con số duy nhất mô tả độ mạnh của một tín hiệu. Có ba con số, và cả ba đều cần:

$$A_{\text{rms}} = \sqrt{\frac{1}{N}\sum_{n=0}^{N-1} x_n^2}\ ; \qquad K = 20\log_{10}\frac{A_{\text{pk}}}{A_{\text{rms}}} \tag{1.4}$$

- **Đỉnh** — giá trị tuyệt đối lớn nhất. Cho biết tín hiệu có vượt ngưỡng và bị méo hay không.
- **Hiệu dụng** — căn bậc hai của trung bình bình phương. Gần với *độ to* cảm nhận hơn nhiều so với đỉnh.
- **Hệ số đỉnh** — khoảng cách giữa hai con số trên, tính bằng dB. Cho biết tín hiệu "nhọn" hay "dày".




(a) Giảm một nửa biên độ là giảm 6 dB




(b) Cùng đỉnh 1,0, năng lượng khác nhau

*Hình 1.11 Ba tín hiệu cùng đỉnh 1,0 nhưng năng lượng và hệ số đỉnh khác hẳn nhau*

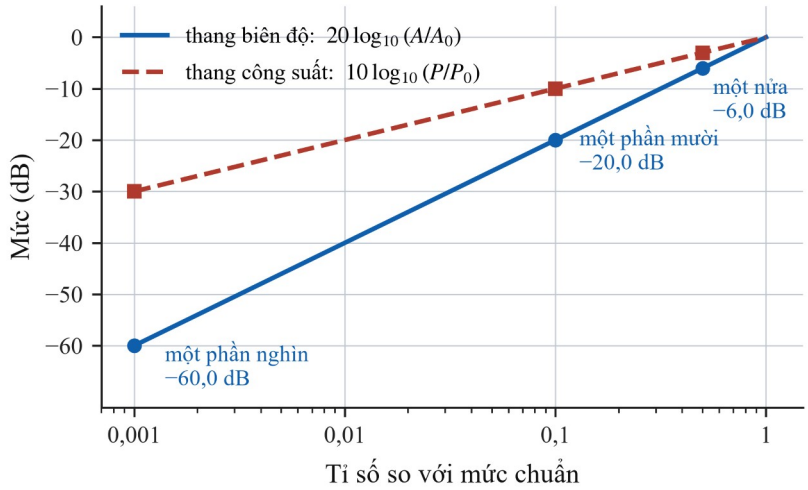
Sóng sin và sóng vuông ở hình trên có cùng đỉnh 1,0, nhưng hiệu dụng của sóng vuông là 1,000 còn của sóng sin chỉ 0,707. Nghĩa là ở cùng mức đỉnh, sóng vuông mang năng lượng lớn hơn và nghe *to hơn* rõ rệt. Hệ số đỉnh của sóng sin là 3,01 dB, của sóng vuông là 0,00 dB.

> **Vì sao điều này quan trọng ngay từ chương 1**
>
> Khi chuẩn hoá âm lượng ở chương 7, câu hỏi đầu tiên sẽ là: chuẩn hoá theo *đỉnh* hay theo *hiệu dụng*? Hai cách cho kết quả rất khác nhau.
>
> Khi học bộ nén khoảng động ở chương 8, việc bộ nén làm chính là *giảm hệ số đỉnh* — kéo đỉnh xuống gần hiệu dụng để tổng thể nghe to hơn mà không vượt ngưỡng.
>
> Hai khái niệm này quay lại nhiều lần, nên đáng bỏ thời gian ở đây.

### 1.3.3 Thang decibel

Tai người cảm nhận độ to theo *tỉ lệ*, nên mọi mức âm thanh đều được đo bằng thang lô-ga-rít gọi là decibel. Đây là chỗ có một điểm dễ nhầm mà gần như ai học cũng gặp một lần: decibel có *hai* công thức.

$$L_A = 20\log_{10}\frac{A}{A_0}; \qquad L_P = 10\log_{10}\frac{P}{P_0} \quad \text{[dB]} \tag{1.5}$$




*Hình 1.12 Hai thang decibel: hệ số 20 cho biên độ, hệ số 10 cho công suất*

*Bảng 1.4 Vài mốc decibel cần thuộc*

| Thay đổi | Theo biên độ | Theo công suất | Cảm nhận |
|---|---|---|---|
| Gấp đôi | +6,02 dB | +3,01 dB | to hơn rõ, nhưng chưa phải "gấp đôi độ to" |
| Một nửa | −6,02 dB | −3,01 dB | nhỏ đi rõ rệt |
| Một phần mười | −20 dB | −10 dB | rất nhỏ |
| Một phần nghìn | −60 dB | −30 dB | gần như im |

**Quy tắc chọn hệ số — không bao giờ sai nữa**

Đang xử lý *dãy mẫu* âm thanh (biên độ, điện áp) → hệ số 20.

Đang xử lý *công suất* hoặc *năng lượng* (kể cả trung bình bình phương dùng cho tỉ số tín hiệu trên nhiễu) → hệ số 10.

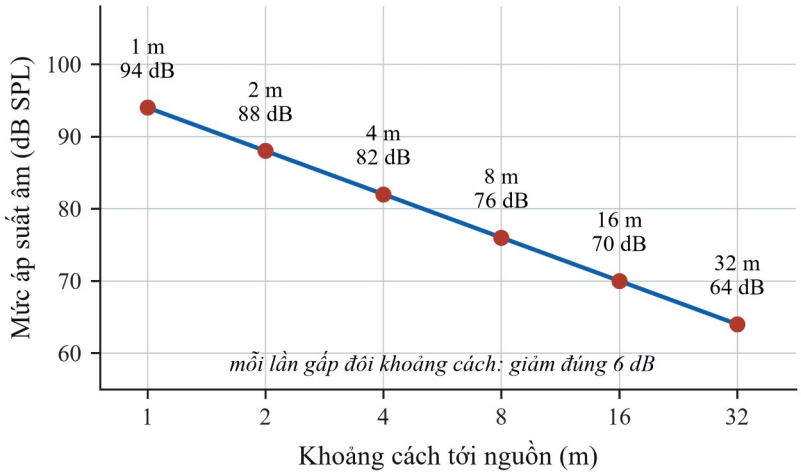
Trong toàn bộ cuốn sách này, chín trên mười lần sẽ dùng hệ số 20 — vì chúng ta gần như luôn làm việc trực tiếp với dãy mẫu.

Nhớ mốc: gấp đôi biên độ là +6 dB, gấp đôi công suất là +3 dB. Thấy con số 6 hay 3 là biết đang ở thang nào.

### 1.3.4 Một hệ quả dùng hằng ngày: quy tắc 6 dB

Với một nguồn âm nhỏ phát ra mọi hướng, năng lượng trải đều lên mặt cầu ngày càng lớn. Kết quả là mức áp suất âm giảm theo khoảng cách theo một quy luật rất gọn:

$$L_2 = L_1 - 20\log_{10}\frac{d_2}{d_1} \tag{1.6}$$




*Hình 1.13 Mỗi lần gấp đôi khoảng cách thì mức giảm đúng 6 dB*

Từ 1 m ra 2 m mất 6 dB. Từ 2 m ra 4 m lại mất đúng chừng ấy. Từ 1 m ra 16 m mất 24 dB.

**Cách dùng quy tắc này ở hiện trường**

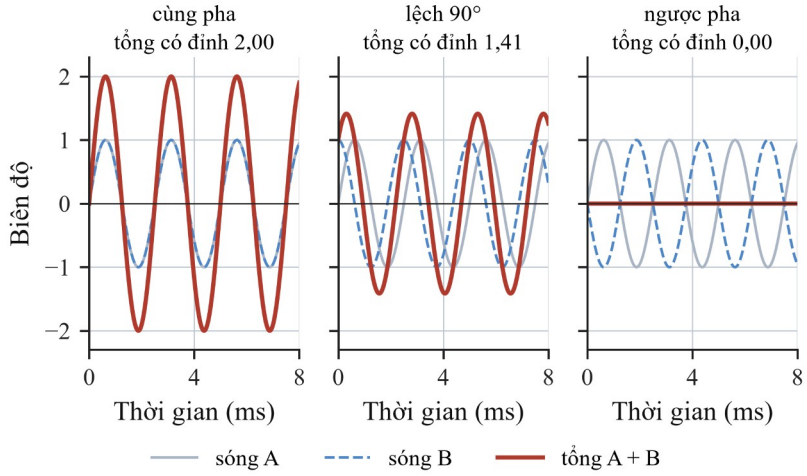
Muốn tín hiệu mạnh hơn 6 dB mà không đụng tới thiết bị: đưa micro lại gần nguồn còn một nửa khoảng cách.

Cái được không chỉ là 6 dB tín hiệu. Nhiễu nền của phòng thì gần như không đổi theo khoảng cách, nên tỉ số tín hiệu trên nhiễu cũng tốt lên đúng 6 dB.

Đây là cách cải thiện chất lượng thu rẻ nhất và hiệu quả nhất — rẻ hơn mọi thiết bị khử nhiễu mà chương 7 sẽ học.

### 1.3.5 Pha — đại lượng vô hình nhưng quyết định

Pha cho biết sóng bắt đầu từ đâu trong chu kỳ của nó. Nghe một sóng sin đơn độc thì không phân biệt được pha — tai không nhạy với nó. Nhưng khi *cộng* hai sóng lại thì pha quyết định tất cả.




*Hình 1.14 Hai sóng giống hệt nhau, chỉ khác pha: tổng có thể gấp đôi hoặc biến mất*

Hai sóng cùng pha cộng lại cho đỉnh 2,00 — gấp đôi. Hai sóng ngược pha cho đỉnh 0,000000 — triệt tiêu hoàn toàn. Con số thứ hai không phải làm tròn: nó thật sự bằng không.

**Hai chỗ gặp pha trong nghề**

*Lỗi đảo pha* khi thu bằng hai micro. Nếu một micro bị nối ngược cực, hoặc đặt ở hai phía của một nguồn, thì trộn hai kênh lại sẽ mất phần lớn âm trầm. Triệu chứng: nghe từng kênh thì bình thường, trộn lại thì mỏng đi.

*Tai nghe chống ồn* hoạt động đúng bằng nguyên lý này: micro ngoài thu tiếng ồn, mạch tạo ra sóng ngược pha, hai sóng triệt tiêu nhau ngay tại màng nhĩ.

Từ chương 9, mọi hiệu ứng trễ điều biến — flanger, chorus — đều dựa trên việc cộng một bản sao lệch pha vào tín hiệu gốc.

## 1.4 Phổ hài và âm sắc

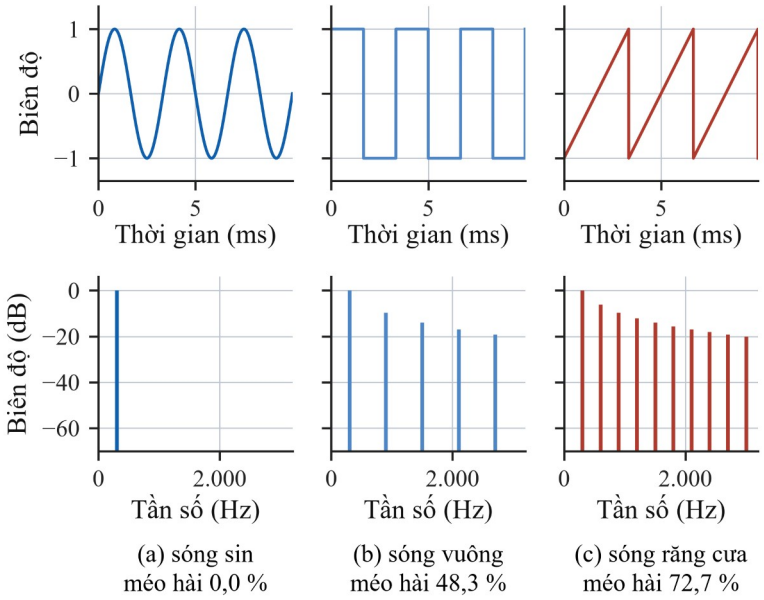
### 1.4.1 Vì sao cùng một nốt mà mỗi nhạc cụ nghe khác nhau

Đàn piano và sáo cùng chơi nốt La4. Cả hai đều cho tần số cơ bản 440 Hz, cùng cao độ, có thể chỉnh cùng độ to. Vậy mà không ai nhầm hai âm ấy với nhau. Cái khác nằm ở *âm sắc*, và âm sắc có một giải thích vật lý rất rõ ràng.

Một vật thật khi dao động không chỉ rung ở một tần số. Nó rung ở tần số cơ bản, đồng thời rung ở các bội số của tần số ấy — gọi là các *hài*. Âm nghe được là tổng của tất cả:

$$x(t) = \sum_{k=1}^{K} A_k \sin(2\pi k f_0 t + \varphi_k) \tag{1.7}$$

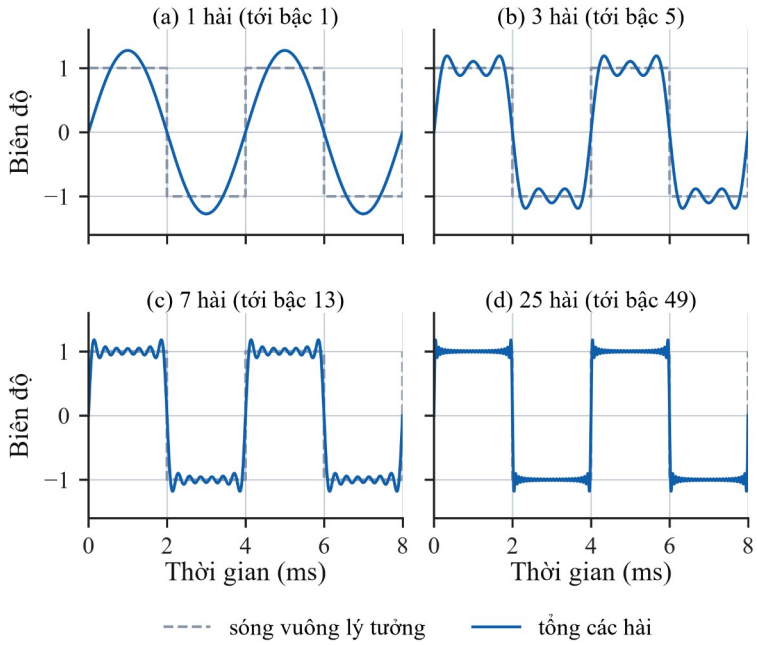
trong đó $f_0$ là tần số cơ bản, k là bậc hài, và $A_k$ là biên độ của hài bậc k. Cao độ do $f_0$ quyết định; âm sắc do *tỉ lệ* các $A_k$ quyết định.




*Hình 1.15 Ba dạng sóng cùng tần số cơ bản 300 Hz và ba phổ tương ứng*

Ba dạng sóng ở hình trên cùng cho cao độ 300 Hz nhưng âm sắc khác hẳn. Hàng dưới cho biết vì sao: sóng sin chỉ có một vạch; sóng vuông có các vạch bậc lẻ; sóng răng cưa có đủ mọi bậc. Độ méo hài đo được lần lượt là 0,00 %, 41,4 % và 72,7 %.

### 1.4.2 Cộng đủ nhiều sóng sin thì ra được mọi dạng sóng

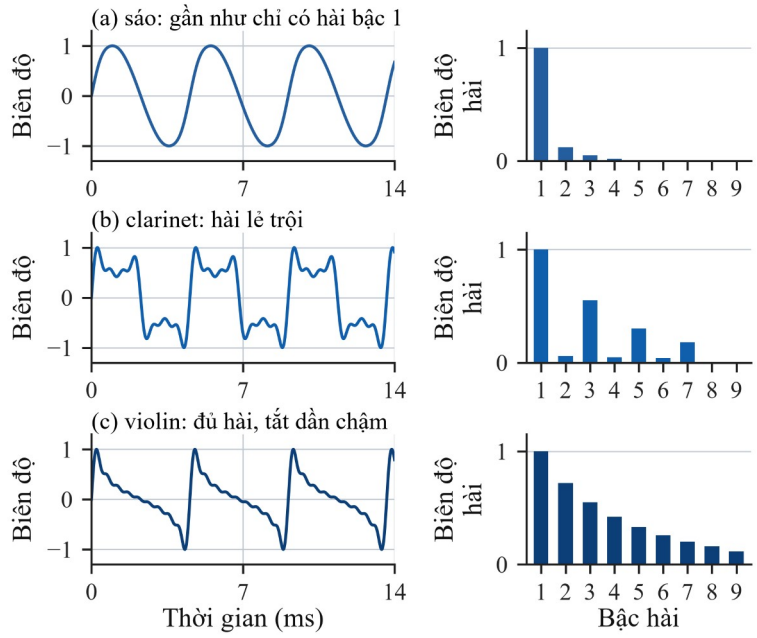



*Hình 1.16 Cộng dần các hài lẻ: từ một sóng sin tới sóng vuông*

Đây là một trong những ý tưởng có sức nặng nhất của kỹ thuật tín hiệu, và nó sẽ được học kỹ ở chương 4 dưới cái tên *biến đổi fourier*. Ở đây chỉ cần thấy điều này: dạng sóng phức tạp nào cũng phân tích được thành tổng các sóng sin, và ngược lại.

Chú ý phần gợn sóng còn lại ở hai mép, ngay cả khi đã cộng hai mươi lăm hài. Đó là hiện tượng Gibbs, và nó không bao giờ mất hẳn dù cộng bao nhiêu hài đi nữa. Hiện tượng này sẽ gặp lại ở chương 4 khi học bộ lọc số, dưới dạng gợn sóng ở dải chắn.

### 1.4.3 Cùng một nốt, ba âm sắc




*Hình 1.17 Ba âm sắc khác nhau dựng từ cùng một tần số cơ bản, chỉ khác tỉ lệ các hài*

Ba dạng sóng ở cột trái được dựng bằng cách cộng các hài với ba bộ biên độ khác nhau ở cột phải. Cả ba đều là nốt La ở 220 Hz — cùng cao độ, cùng độ to. Cái khác duy nhất là tỉ lệ biên độ các hài, và chỉ thế thôi đã đủ để tai nhận ra ba nhạc cụ khác nhau.

*Bảng 1.5 Đặc điểm phổ hài của vài nhóm nhạc cụ*

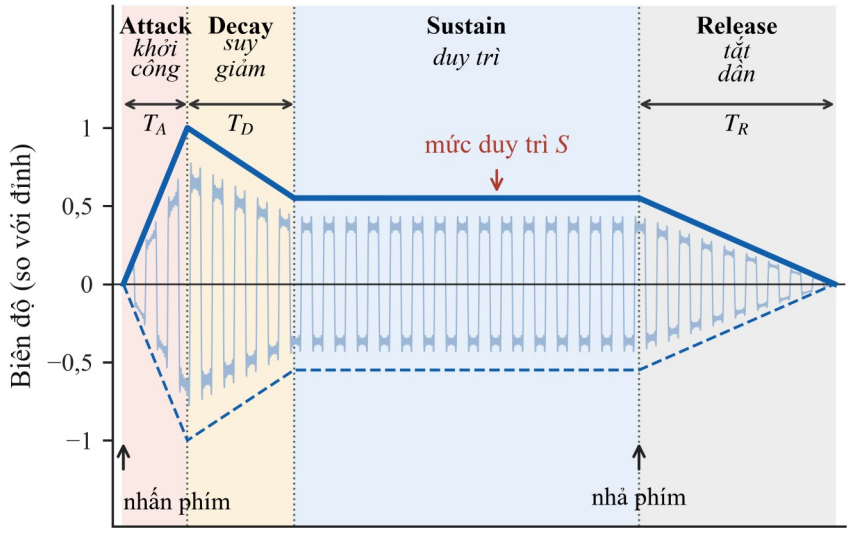
| Nhóm | Đặc điểm phổ hài | Nghe ra sao |
|---|---|---|
| Sáo, ống thổi hở | Hài bậc 1 áp đảo, các hài trên rất yếu | Trong, mềm, gần với sóng sin thuần |
| Clarinet, ống bịt một đầu | Hài *lẻ* trội hẳn, hài chẵn rất yếu | Rỗng, hơi "mũi", đặc trưng dễ nhận |
| Violin, đàn dây kéo | Đủ cả hài chẵn lẫn lẻ, tắt dần chậm | Sáng, dày, nhiều chi tiết |
| Trống, chuông | Các thành phần *không* phải bội số nguyên của $f_0$ | Không có cao độ rõ ràng |

Dòng cuối bảng đáng chú ý: khi các thành phần không phải bội số nguyên của tần số cơ bản thì tai không quy được về một cao độ. Đó là lý do gõ trống không "ra nốt" như gảy đàn.

### 1.4.4 Đường bao biên độ ADSR — âm sắc còn nằm ở thời gian

Các mục trên giải thích âm sắc bằng tỉ lệ các hài, tức bằng hình dạng của phổ ở một khoảnh khắc. Nhưng một nốt nhạc thật không phải một khoảnh khắc: nó có lúc bắt đầu, có đoạn kéo dài, và có lúc tắt. Cách biên độ của nốt biến đổi từ lúc bắt đầu tới lúc tắt gọi là *đường bao biên độ* (amplitude envelope). Cùng một dạng sóng và cùng một phổ hài, nhân với hai đường bao khác nhau, cho ra hai âm thanh nghe khác hẳn nhau: một tiếng gõ tắt nhanh và một tiếng kéo dài đều đặn.

Mô hình đơn giản và được dùng nhiều nhất để mô tả đường bao là *ADSR*, viết tắt của bốn giai đoạn nối tiếp nhau. Hình 1.18 vẽ bốn giai đoạn ấy cho một nốt được giữ phím rồi nhả, và Bảng 1.6 ghi ý nghĩa của từng tham số.




*Hình 1.18 Đường bao biên độ ADSR của một nốt: bốn giai đoạn và bốn tham số*

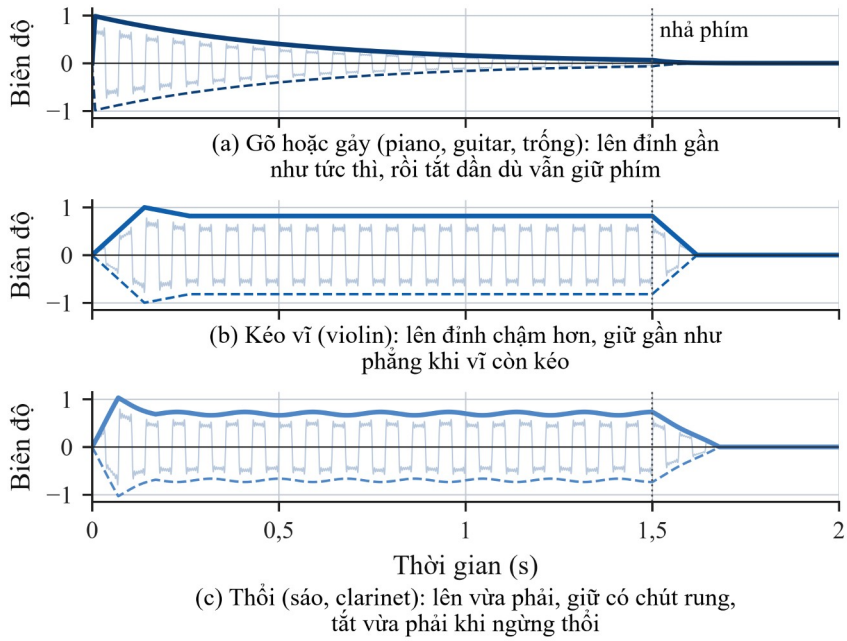
*Bảng 1.6 Bốn giai đoạn của đường bao ADSR*

| Giai đoạn | Tham số | Bắt đầu và kết thúc | Ý nghĩa nghe được |
|---|---|---|---|
| **Attack** — khởi công | $T_A$ (thời gian) | Từ lúc nhấn phím, biên độ tăng từ 0 lên đỉnh. | Nốt bắt đầu gắt hay mềm. Càng ngắn nghe càng "bật" và sắc. |
| **Decay** — suy giảm | $T_D$ (thời gian) | Từ đỉnh, biên độ hạ xuống mức duy trì. | Phần "nảy" sau cú khởi đầu; nhiều nhạc cụ thật có phần này rõ. |
| **Sustain** — duy trì | $S$ (mức, không phải thời gian) | Biên độ giữ nguyên chừng nào phím còn được giữ. | Nốt kéo dài ổn định hay tắt dần. Đặt $S = 0$ thì nốt chỉ còn khởi đầu rồi tắt. |
| **Release** — tắt dần | $T_R$ (thời gian) | Từ lúc nhả phím, biên độ hạ từ mức hiện có về 0. | Nốt dứt khoát hay còn ngân. Càng dài càng nghe như có vang. |

Một chỗ rất hay nhầm: ba tham số $T_A$, $T_D$ và $T_R$ là *thời gian* (đo bằng giây hoặc mili giây), còn $S$ là một *mức biên độ* (thường tính theo tỉ lệ với đỉnh, từ 0 đến 1). Vì vậy giai đoạn duy trì không có độ dài riêng: nó kéo dài đúng bằng khoảng thời gian phím còn được giữ.

> **ADSR chỉ là mô hình**
> Nhạc cụ thật hiếm khi có đường bao gồm bốn đoạn thẳng. Đàn piano không có giai đoạn duy trì thật sự: dây tắt dần suốt thời gian phím còn được giữ. Đàn violin có thể tăng giảm biên độ theo tay vĩ. Bộ tổng hợp âm thanh dùng ADSR vì bốn tham số này đủ để gần đúng phần lớn nốt nhạc và đủ ít để người dùng vặn bằng bốn núm.

Mô hình ADSR cho thấy ba nhóm nhạc cụ có cách "sống" rất khác nhau, dù có thể phát cùng một nốt với cùng một phổ hài. Hình 1.19 vẽ ba kiểu đường bao điển hình trên cùng một dạng sóng.




*Hình 1.19 Cùng một dạng sóng, ba đường bao khác nhau: gõ hoặc gảy, kéo vĩ, thổi (tần số sóng được hạ thấp để nhìn rõ)*

*Bảng 1.7 Đường bao điển hình của vài nhóm nhạc cụ (mức độ gần đúng, mang tính minh hoạ)*

| Nhóm | Attack | Decay và Sustain | Release |
|---|---|---|---|
| Piano, guitar gảy, trống | Rất ngắn, cỡ vài mili giây | Tắt dần liên tục, gần như không có đoạn phẳng | Ngắn nếu có bộ giảm âm (piano nhả phím thì dập), dài nếu để ngân |
| Đàn dây kéo vĩ (violin, cello) | Vừa, cỡ vài chục tới vài trăm mili giây | Giữ gần như phẳng khi vĩ còn kéo | Ngắn, khi nhấc vĩ |
| Kèn sáo (flute, clarinet, oboe) | Vừa, thường kèm tiếng hơi ở đầu nốt | Giữ phẳng, có chút rung biên độ | Ngắn tới vừa |
| Đàn organ ống | Mềm, hơi chậm | Phẳng hoàn toàn, mức duy trì cao | Ngắn |

**Vì sao đường bao quan trọng với tai?** Có ba lý do.

– **Phần khởi công mang nhiều thông tin nhận dạng nguồn âm.** Lúc nốt vừa bắt đầu, nhạc cụ thường phát ra nhiều hài cao và cả thành phần ồn ào chưa ổn định, rồi mới dần về trạng thái duy trì. Các thí nghiệm nghe cắt bỏ đoạn đầu của nốt cho thấy người nghe khó nhận ra nhạc cụ hơn nhiều khi chỉ còn đoạn giữa, dù phổ hài của đoạn giữa không đổi. Vì thế đường bao là một nửa của âm sắc, chứ không phải chi tiết phụ.

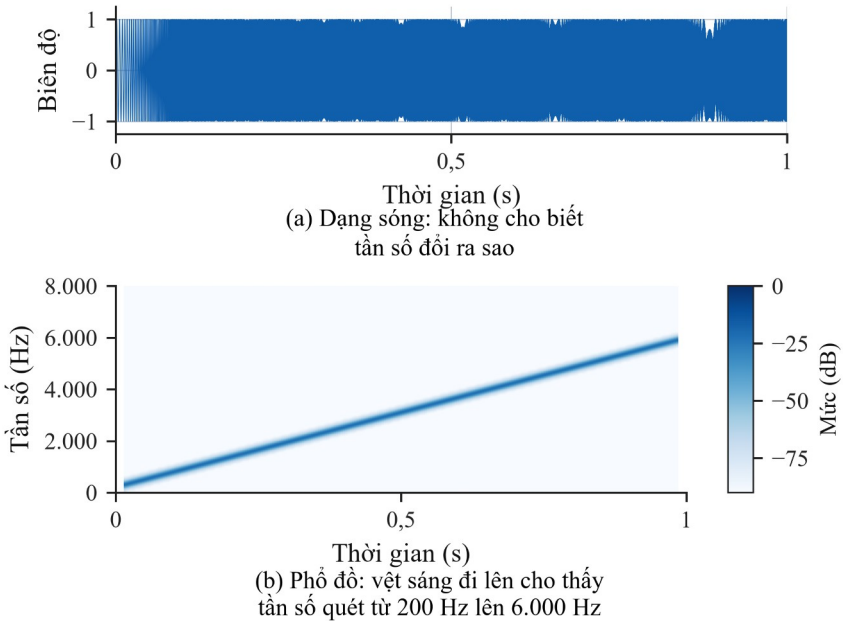
- **Phổ thay đổi theo thời gian.** Phổ hài ở mục 1.4.1 mô tả trạng thái duy trì. Ở giai đoạn khởi công và tắt dần, tỉ lệ các hài còn đổi theo từng thời điểm; muốn nhìn thấy điều đó phải dùng phổ đồ ở mục 1.4.5.
- **Đường bao quyết định cảm giác về nhịp và khoảng cách.** Nốt có attack ngắn nghe gần và rõ nhịp; attack dài nghe mềm và lùi xa. Release dài nghe như có vang của căn phòng.

**Đường bao trong lập trình.** Cuốn sách này gặp lại đường bao ở bốn chỗ, và mỗi chỗ dùng nó cho một việc khác. Ở chương 6, đường bao được dùng để làm fade vào, fade ra và chữa tiếng "tách": khi một nốt bắt đầu hoặc kết thúc đột ngột (attack hoặc release bằng 0), đồ thị nhảy bậc và tai nghe thấy tiếng click. Ở chương 8, "attack" và "release" xuất hiện trong bộ nén và bộ cổng, nhưng ở đó chúng là thời gian phản ứng của bộ xử lý, không phải đường bao của nguồn âm. Ở chương 11 và 12, Web Audio lập lịch đường bao cho một tham số (ví dụ âm lượng) bằng các lệnh hẹn giờ trên AudioParam. Hiểu ADSR ở đây là để nhận ra cùng một ý tưởng mang bốn tên gọi khác nhau.

> **Thử nghe ngay**
> Mở Audacity, thu một tiếng gõ ngón tay lên mặt bàn và một nốt ngân dài bằng giọng hát "aaa". Phóng to để thấy hai đường bao. Rồi cắt bỏ khoảng 100 mili giây đầu của mỗi đoạn và nghe lại. Tiếng gõ gần như mất cá tính, còn nốt hát thì vẫn nhận ra là giọng người nhưng nghe "cụt" ở đầu. Ghi lại ba nhận xét; chúng sẽ dùng lại khi làm bài tập về fade ở chương 6.

### 1.4.5 Nhìn tần số thay đổi theo thời gian

Phổ ở các hình trước được tính cho cả đoạn tín hiệu, nên nó cho biết đoạn ấy có những tần số nào, nhưng không cho biết chúng xuất hiện *lúc nào*. Với âm thanh thật — một câu nói, một câu nhạc — thì tần số thay đổi liên tục, và chúng ta cần một cách nhìn khác.




*Hình 1.20 Phổ đồ của một tín hiệu quét tần số: dạng sóng không cho biết gì, phổ đồ thì nói hết*

**Cách đọc một phổ đồ**

Trục ngang là *thời gian*, trục dọc là *tần số*, độ sáng là *độ mạnh*.

Một vạch ngang sáng nghĩa là có một tần số ổn định kéo dài. Một vệt chéo đi lên nghĩa là tần số đang tăng dần.

Đây là công cụ nhìn quan trọng nhất của cả cuốn sách. Chương 4 sẽ học cách nó được tính; từ giờ tới đó chỉ cần biết *đọc*.

Ba cách nhìn vừa đi qua sẽ theo suốt cuốn sách, và mỗi cách trả lời một câu hỏi khác nhau. *Dạng sóng* trả lời "biên độ thay đổi thế nào theo thời gian" — đó là cách nhìn của chương 2 và chương 3, khi bàn về lấy mẫu và lượng tử hoá. *Phổ* trả lời "đoạn này gồm những tần số nào" — cách nhìn của chương 4 và của toàn bộ phần lọc ở chương 5. *Phổ đồ* trả lời "tần số nào xuất hiện lúc nào" — cách nhìn không thể thiếu khi làm việc với tiếng nói và âm nhạc thật, từ chương 9 trở đi. Chọn nhầm cách nhìn thì thứ đang tìm sẽ vô hình: một tiếng rít 8 kHz lẫn trong bản thu gần như không thấy trên dạng sóng, nhưng hiện ra thành một vạch ngang rõ mồn một trên phổ đồ.

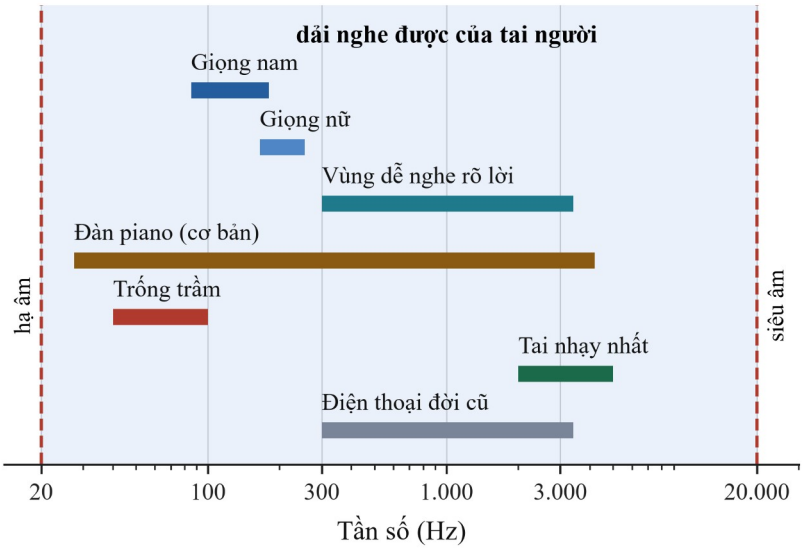
Một việc nên làm ngay sau khi đọc xong chương: mở Audacity, thu vài giây giọng nói của chính mình, rồi bật chế độ Spectrogram và so với chế độ Waveform. Nguyên âm hiện ra thành những vạch ngang song song — đó chính là các hài vừa học ở mục 1.4.2. Phụ âm gió như "s" hay "x" hiện ra thành một mảng sáng tản mạn ở vùng tần số cao, không có vạch nào cả, vì chúng gần với nhiễu hơn là với âm có cao độ. Nhìn thấy sự khác nhau ấy một lần bằng đọc mười lần.

## 1.5 Tai người nghe như thế nào

Ba mục trước nói về âm thanh như một hiện tượng vật lý. Mục này nói về cái *nhận* nó. Đây là phần hay bị coi nhẹ nhất trong một môn kỹ thuật, và cũng là

phần có ích nhất — vì mọi quyết định kỹ thuật về sau đều nhắm tới một cái tai, chứ không nhắm tới một cái máy đo.

### 1.5.1 Dải nghe được

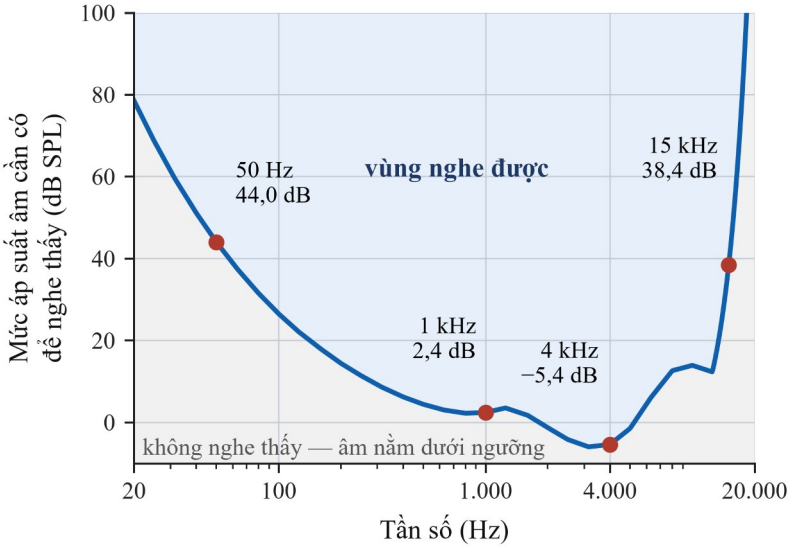



*Hình 1.21 Dải nghe được của tai người và chỗ đứng của các nguồn âm quen thuộc*

Dải nghe được thường ghi là 20 Hz tới 20.000 Hz. Hai con số này là giá trị của người trẻ tai tốt, và giới hạn trên tụt dần theo tuổi — tới tuổi bốn mươi thường chỉ còn khoảng 14 đến 16 kHz. Âm dưới 20 Hz gọi là hạ âm, trên 20 kHz gọi là siêu âm; cả hai đều tồn tại nhưng tai không nhận được.

Dòng gần cuối hình đáng để ý: điện thoại đời cũ chỉ truyền dải 300 tới 3.400 Hz, tức chưa tới một phần năm dải nghe được, vậy mà vẫn nghe rõ lời. Lý do là toàn bộ thông tin phân biệt phụ âm nằm gọn trong dải ấy. Đây là ví dụ đầu tiên của một ý sẽ trở lại nhiều lần: không phải phần nào của tín hiệu cũng quan trọng như nhau.

### 1.5.2 Tai không nhạy đều theo tần số




*Hình 1.22 Ngưỡng nghe tuyệt đối — mức cần có để nghe thấy, theo từng tần số*

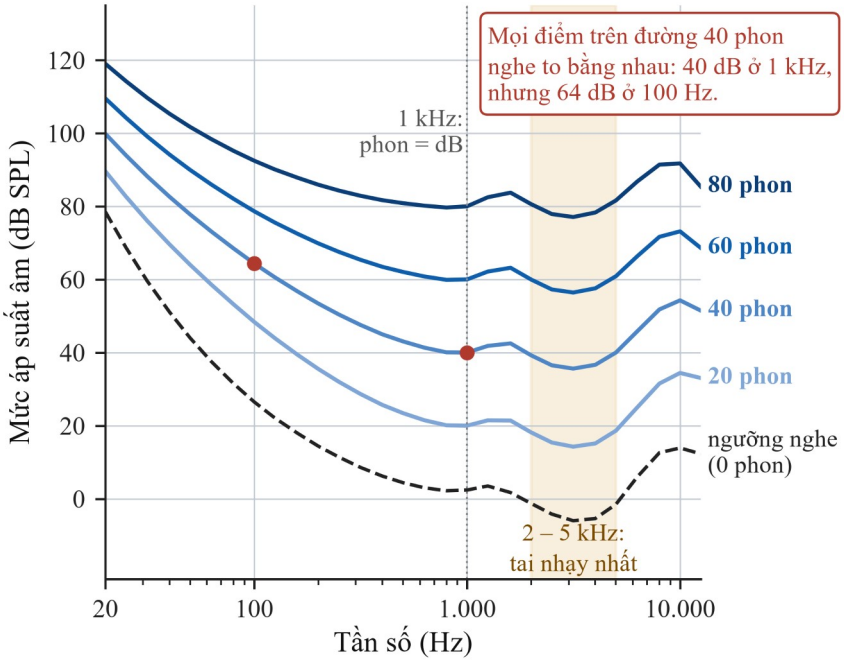
Đường cong này (dữ liệu ISO 226:2003) cho biết một âm phải mạnh bao nhiêu thì mới nghe thấy. Ở 4.000 Hz chỉ cần khoảng −5 dB. Ở 1.000 Hz cần khoảng 2 dB. Còn ở 50 Hz thì phải tới 44 dB — tức mạnh hơn khoảng 300 lần về biên độ so với ở 4 kHz, mà vẫn chỉ vừa đủ nghe.

> **Ba hệ quả thực hành của đường cong này**
>
> Vùng 2 – 5 kHz là vùng tai nhạy nhất, cũng là vùng chứa phần lớn thông tin phân biệt phụ âm. Đụng vào vùng này khi chỉnh âm sắc thì hậu quả rõ hơn hẳn so với đụng vào vùng khác.
>
> Âm trầm phải phát mạnh hơn nhiều mới nghe thấy. Đó là lý do loa trầm cần công suất lớn, và cũng là lý do nghe nhạc nhỏ thì thấy "mất trầm".
>
> Đây là *điểm mù thứ nhất* của tai. Bộ mã hoá có tổn hao khai thác nó: âm nằm dưới đường này thì không cần truyền đi.

### *Đường đẳng âm và đơn vị phon*

Hình 1.22 chỉ là một đường: mức tối thiểu để nghe thấy. Nhưng độ to cảm nhận không chỉ phụ thuộc vào mức áp suất âm, mà còn phụ thuộc vào tần số, và sự phụ thuộc ấy lại thay đổi theo mức nghe. Để đo điều này, Fletcher và Munson (1933) cho người nghe so sánh các âm: lấy một âm chuẩn 1 kHz ở một mức nào đó, rồi chỉnh âm ở tần số khác cho tới khi người nghe thấy hai âm *to bằng nhau*. Nối các điểm ấy lại ta được một *đường đẳng âm* (equal-loudness contour). Kết quả của nhiều nhóm nghiên cứu về sau đã được chuẩn hoá trong ISO 226 (bản 2003), và các số liệu trong mục này lấy từ chuẩn ấy.

Mỗi đường đẳng âm mang một số gọi là *phon*. Một âm có độ to $N$ phon là âm nghe to bằng một âm 1 kHz ở mức $N$ dB SPL. Vì thế, ở 1 kHz thì phon và dB trùng nhau; ở các tần số khác thì chúng khác nhau. Ví dụ ở Hình 1.23, đường 40 phon đi qua 40 dB ở 1 kHz, nhưng ở 100 Hz phải tới 64 dB, và ở 50 Hz phải tới 78 dB mới nghe to bằng.




*Hình 1.23 Đường đẳng âm theo ISO 226:2003: mỗi đường nối các âm nghe to bằng nhau, đánh số bằng phon*

*Bảng 1.8 Mức áp suất âm (dB SPL) cần có để nghe to bằng nhau, theo ISO 226:2003*

| Tần số | 20 phon | 40 phon | 60 phon | 80 phon |
|---|---|---|---|---|
| 50 Hz | 64 | 78 | 90 | 102 |
| 100 Hz | 48 | 64 | 79 | 92 |
| 1.000 Hz | 20 | 40 | 60 | 80 |
| 4.000 Hz | 15 | 37 | 58 | 78 |
| 8.000 Hz | 32 | 52 | 72 | 91 |
| Chênh lệch 50 Hz so với 1 kHz | 44 dB | 38 dB | 30 dB | 22 dB |

Từ hình và bảng có thể rút ra hai nhận xét. Thứ nhất, các đường thấp nhất ở vùng 2 – 5 kHz: tai nhạy nhất ở đó, một phần vì ống tai cộng hưởng khoảng vài kilôhéc, giống một ống một đầu kín. Thứ hai, và quan trọng hơn trong thực hành, các đường *phẳng dần khi to lên*: để một âm 50 Hz nghe to bằng âm 1 kHz thì ở 20 phon phải mạnh hơn 44 dB, nhưng ở 80 phon chỉ cần mạnh hơn khoảng 22 dB.

**Vì sao nghe nhỏ thì thấy "mất trầm"**

Khi giảm âm lượng, mọi tần số cùng giảm một lượng dB như nhau, nhưng tai thì không giảm độ to đều như thế: các đường đẳng âm càng thấp càng dốc ở phía trầm. Bản nhạc được phối ở mức to nghe sẽ thiếu trầm khi mở nhỏ. Nút "loudness" trên nhiều ampli là một bộ cân bằng tự nâng trầm và cao khi âm lượng hạ, để bù lại hiệu ứng này. Người làm hậu kỳ vì thế phải nghe kiểm tra ở nhiều mức, chứ không chỉ ở một mức.

> **Liên hệ với dBA**
> Máy đo tiếng ồn thường hiển thị mức theo trọng số A (dBA). Bộ lọc trọng số A xấp xỉ dạng của đường đẳng âm mức thấp (cỡ 40 phon): nó hạ phần trầm và phần cao thật cao, vì tai kém nhạy ở đó. Nhờ vậy một con số dBA gần hơn với cảm giác ồn so với mức dB SPL không trọng số. Cách đo ghi trong IEC 61672, đã nêu ở danh mục tài liệu của chương này.

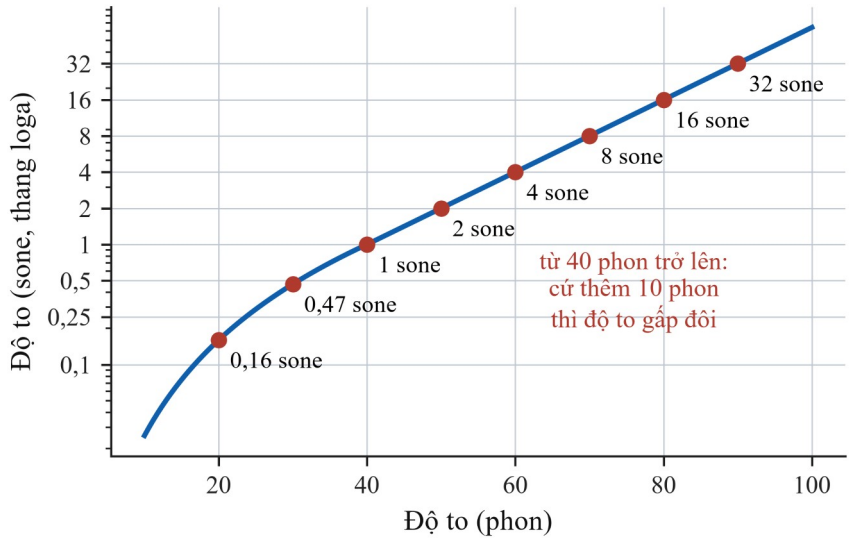
### *Sone — đo độ to bằng đơn vị tỉ lệ*

Phon cho biết hai âm có to bằng nhau hay không, nhưng không cho biết một âm to gấp mấy lần âm kia: âm 60 phon không phải "to gấp ba" âm 20 phon. Thang *sone* (Stevens, 1936) sinh ra để trả lời câu hỏi đó. Quy ước: **1 sone** là độ to của âm 1 kHz ở 40 phon, và độ to cảm nhận gấp đôi thì số sone cũng gấp đôi. Nói cách khác, sone là thang tỉ lệ còn phon là thang so sánh.

Từ 40 phon trở lên, thí nghiệm cho thấy cứ **tăng 10 phon thì độ to cảm nhận gấp đôi**, nên:

$$N = 2^{(P-40)/10} \qquad [\text{sone}] \tag{1.8}$$

trong đó P là độ to tính bằng phon ($P \geq 40$) và N là độ to tính bằng sone. Dưới 40 phon, độ to tăng chậm hơn nhiều so với quy tắc này và thường được xấp xỉ bằng $N \approx (P/40)^{2{,}642}$. Hình 1.24 vẽ quan hệ đó, và Bảng 1.9 ghi vài mốc kèm ví dụ gần đúng cho dễ hình dung.




*Hình 1.24 Quan hệ giữa phon và sone: từ 40 phon trở lên, thêm 10 phon thì độ to gấp đôi (thang sone là thang loga)*

*Bảng 1.9 Vài mốc phon và sone (ví dụ chỉ mang tính minh hoạ)*

| Độ to | Số sone | Ví dụ gần đúng |
|---|---|---|
| 20 phon | 0,16 sone | Gần ngưỡng nghe ở tần số giữa |
| 30 phon | 0,47 sone | Phòng ngủ rất yên tĩnh |
| 40 phon | 1 sone | Phòng yên tĩnh, thư viện |
| 50 phon | 2 sone | Nói chuyện nhỏ, văn phòng yên |
| 60 phon | 4 sone | Trò chuyện bình thường |
| 70 phon | 8 sone | Phố đông, máy hút bụi ở xa |
| 80 phon | 16 sone | Phố ồn, nhạc nghe to |
| 90 phon | 32 sone | Rất to; nghe lâu có hại |

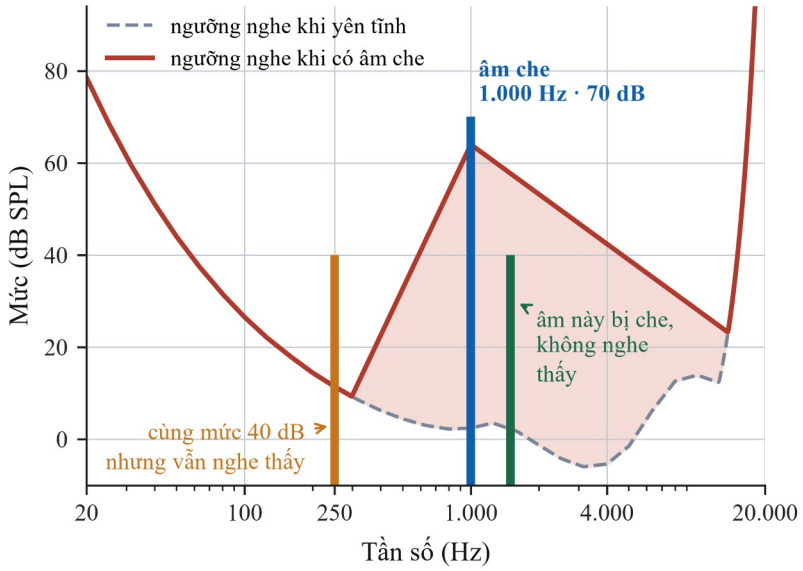
**Ba cách nói "gấp đôi" — đừng nhầm**

Biên độ gấp đôi tương ứng +6 dB (mục 1.3.4). Công suất gấp đôi tương ứng +3 dB. Nhưng *độ to cảm nhận gấp đôi* tương ứng khoảng +10 dB ở vùng giữa và từ 40 phon trở lên. Tức là phải tăng biên độ chừng 3 lần (khoảng 10 dB) thì tai mới thấy "to gấp đôi". Đây là lý do một bộ khuếch đại công suất gấp đôi chỉ làm người nghe thấy to hơn một chút.

Cần nhớ rằng độ to là đại lượng *chủ quan*, đo bằng thí nghiệm nghe, nên giá trị thật khác nhau đôi chút giữa người này với người khác, và quy tắc "10 phon gấp đôi" chỉ là một xấp xỉ tốt. Với âm phức tạp như nhạc hay giọng nói, độ to còn phụ thuộc vào cả dải tần và thời lượng của âm; chuẩn ISO 532 tính độ to sone cho các âm như vậy. Trong thực hành xử lý tệp âm thanh, chương 7 sẽ dùng một thước đo khác, LUFS, được thiết kế để phản ánh độ to cảm nhận của cả một đoạn âm thanh.

### 1.5.3 Hiện tượng che lấp

Điểm mù thứ hai quan trọng hơn nhiều. Khi có một âm mạnh, ngưỡng nghe của cả vùng tần số lân cận bị *nâng lên* — những âm yếu ở gần đó, vốn nghe được khi yên tĩnh, nay không nghe thấy nữa.

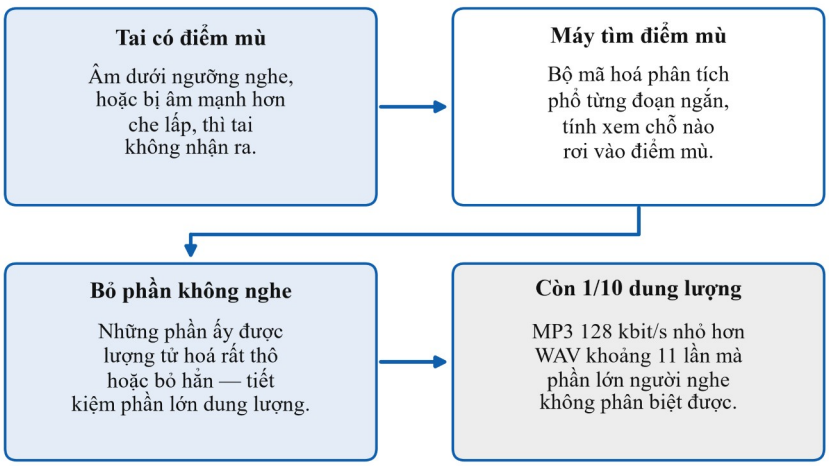



*Hình 1.25 Che lấp theo tần số: một âm mạnh nâng ngưỡng nghe của vùng lân cận*

Hình trên cho thấy một chi tiết quan trọng: búp che lấp *không* đối xứng. Nó dốc đứng về phía tần số thấp và thoải về phía tần số cao. Nghĩa là một âm mạnh che các âm *cao hơn* nó dễ hơn nhiều so với các âm trầm hơn. Hai vạch trong hình có cùng mức 40 dB, nhưng vạch nằm phía trên bị che hoàn toàn còn vạch nằm phía dưới thì vẫn nghe rõ.

Ngoài che lấp theo tần số còn có che lấp theo *thời gian*: một âm mạnh che các âm yếu xuất hiện ngay sau nó chừng vài chục mili giây, và kỳ lạ hơn, che cả những âm xuất hiện ngay *trước* nó vài mili giây. Chương 3 sẽ nói kỹ hơn khi bàn về nén.

### 1.5.4 Vì sao phần này nằm ở chương 1




*Hình 1.26 Từ một hiện tượng sinh lý tới định dạng nén có tổn hao*

Chuỗi ở hình trên là lý do phần thính giác được đặt ngay chương đầu chứ không phải chương cuối. Nén có tổn hao — MP3, AAC, Opus — không phải một kỹ thuật toán học khó hiểu. Nó chỉ là việc tìm ra những phần mà tai không nghe được rồi bỏ đi. Ai hiểu che lấp thì hiểu nén; ai không hiểu thì chương 3 sẽ thành một danh sách định dạng phải học thuộc.

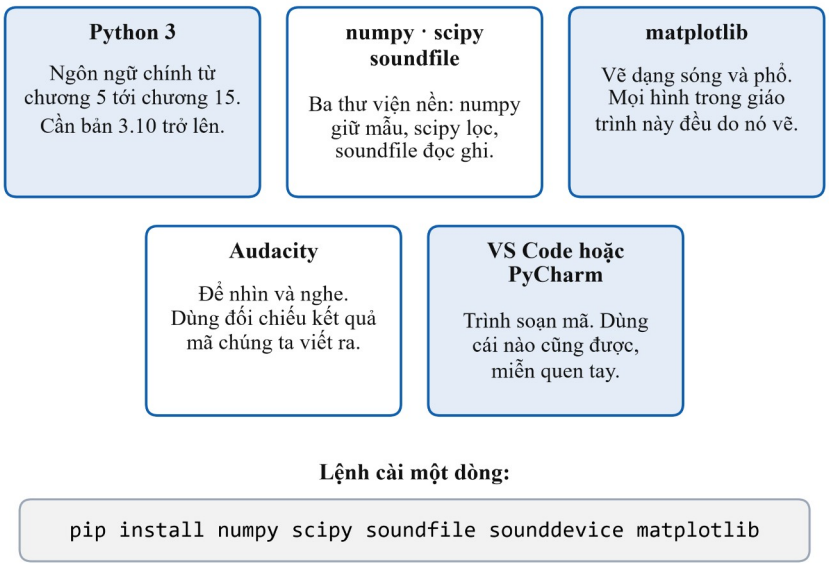
Con số cho thấy sức nặng của việc ấy: một phút âm thanh chất lượng CD chiếm 10,6 MB. Cùng một phút ấy ở MP3 128 kbit/s chỉ còn 0,96 MB, tức nhỏ hơn khoảng 11 lần, mà phần lớn người nghe không phân biệt được. Toàn bộ khoản tiết kiệm ấy đến từ hai điểm mù vừa học.

Có một điều cần nói rõ để tránh hiểu sai về sau. "Tai không nghe được" không phải một tính chất cố định của âm thanh, mà là một kết luận *có điều kiện*. Cùng một thành phần yếu ấy, khi nằm cạnh một âm mạnh thì bị che, nhưng khi âm mạnh tắt đi thì lại nghe rõ. Khi nghe qua loa điện thoại trong quán ồn thì không ai nhận ra, nhưng khi nghe bằng tai nghe tốt trong phòng yên thì nhiều người nhận ra. Vì vậy mọi bộ nén có tổn hao đều phải chọn trước một *giả định* về điều kiện nghe, và chất lượng của nó phụ thuộc vào việc giả định ấy có đúng với người nghe thật hay không.

Hệ quả thực hành quan trọng nhất của điều vừa nói là: đừng nén có tổn hao nhiều lần chồng lên nhau. Mỗi lần nén, bộ mã hoá lại bỏ đi những phần mà nó cho là bị che, dựa trên tín hiệu *đang có* trước mặt nó. Sau lần nén thứ nhất, chính những âm mạnh vốn đang che cho các âm yếu đã bị thay đổi; lần nén thứ hai vì thế tính nhầm, và bỏ đi cả những phần thực sự còn nghe được. Đó là lý do quy tắc nghề nghiệp luôn là: giữ bản gốc không nén để làm việc, chỉ nén ở bước cuối cùng khi xuất bản. Quy tắc ấy sẽ được nhắc lại kèm số liệu cụ thể ở chương 3.

## 1.6 Hướng dẫn thực hành

### 1.6.1 Cài môi trường




*Hình 1.27 Môi trường làm việc dùng chung cho cả mười lăm chương*

Cài một lần, dùng cho cả cuốn sách. Trên máy đã có Python 3.10 trở lên, mở cửa sổ dòng lệnh và chạy:

*Liệt kê 1.1 Cài các thư viện cần thiết*

```
pip install numpy scipy soundfile sounddevice matplotlib
```

Riêng Audacity thì tải và cài như một phần mềm bình thường. Nó không tham gia vào mã, nhưng sẽ dùng suốt cuốn sách để nhìn và nghe kết quả — đối chiếu với những gì mã tự viết làm ra.

### 1.6.2 Đoạn mã kiểm môi trường

Tạo tệp kiem_moi_truong.py với nội dung sau rồi chạy. Đây cũng là đoạn mã đầu tiên của cuốn sách, và nó đã chứa gần hết những gì chương này bàn tới.

*Liệt kê 1.2 kiem_moi_truong.py — sinh một sóng sin, nghe, nhìn và đo*

```
import numpy as np
import matplotlib.pyplot as plt
import soundfile as sf

SR = 44100                    # tần số lấy mẫu, mẫu mỗi giây
F0 = 440.0                    # nốt La4
DUR = 2.0                     # thời lượng, giây

# ── sinh tín hiệu: đúng công thức (1.1) ──
t = np.arange(int(DUR * SR)) / SR          # trục thời gian
x = 0.5 * np.sin(2 * np.pi * F0 * t)       # biên độ 0,5

# ── đo ──
dinh = np.max(np.abs(x))
hieu_dung = np.sqrt(np.mean(x ** 2))
print("đỉnh      =", round(dinh, 4))
print("hiệu dụng =", round(hieu_dung, 4))
print("hệ số đỉnh=", round(20 * np.log10(dinh / hieu_dung), 2), "dB")

# ── ghi ra tệp để nghe bằng Audacity ──
sf.write("sin440.wav", x, SR)

# ── nhìn: dạng sóng và phổ ──
fig, (ax1, ax2) = plt.subplots(1, 2, figsize=(11, 3.5))
ax1.plot(t[:441] * 1000, x[:441])          # 10 ms đầu
ax1.set_xlabel("Thời gian (ms)")

w = np.hanning(len(x))
X = np.abs(np.fft.rfft(x * w)) / np.sum(w) * 2
f = np.fft.rfftfreq(len(x), 1 / SR)
ax2.plot(f, 20 * np.log10(np.maximum(X / np.max(X), 1e-10)))
ax2.set_xlim(0, 2000); ax2.set_ylim(-80, 5)
ax2.set_xlabel("Tần số (Hz)")
plt.tight_layout(); plt.show()

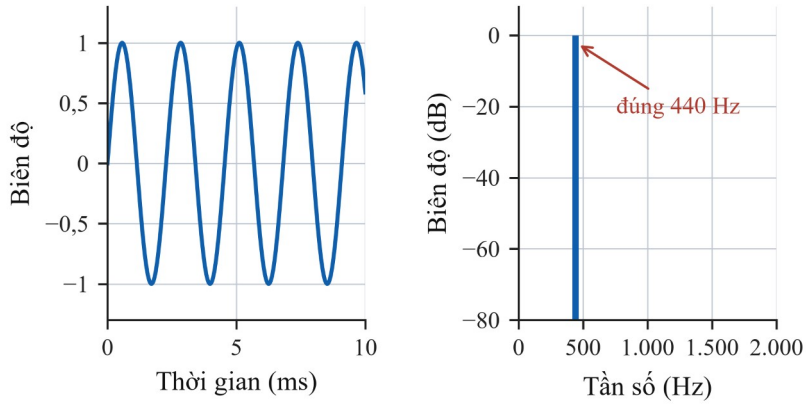
print("đỉnh phổ nằm ở", round(f[np.argmax(X)], 1), "Hz")
```

**Kết quả phải ra**

đỉnh = 0,5000 · hiệu dụng = 0,3536 · hệ số đỉnh = 3,01 dB

đỉnh phổ nằm ở 440,0 Hz

Cửa sổ đồ hoạ hiện ra hai biểu đồ như hình dưới. Nghe tệp sin440.wav thấy một tiếng "u" đều, không đổi cao độ.




(a) Dạng sóng: sóng sin đều đặn (b) Phổ: đúng một vạch

*Hình 1.28 Dạng sóng và phổ phải thấy sau khi chạy đoạn mã kiểm môi trường*

Chú ý con số hệ số đỉnh: nó ra đúng 3,01 dB, khớp với giá trị lý thuyết của sóng sin đã nêu ở mục 1.3.2. Đây là lần đầu trong cuốn sách một con số tính bằng mã khớp với một con số suy ra từ công thức — và đó là cách làm việc sẽ theo suốt mười lăm chương.

### 1.6.3 Ba lỗi hay gặp khi chạy lần đầu

*Bảng 1.10 Triệu chứng, nguyên nhân và cách sửa*

| Triệu chứng | Nguyên nhân thường gặp | Cách sửa |
|---|---|---|
| ModuleNotFoundError: No module named ... | Thư viện chưa cài, hoặc cài vào một bản Python khác với bản đang chạy | Chạy python -m pip install thay vì pip install, để chắc chắn cài đúng bản đang dùng |
| Cửa sổ đồ hoạ không hiện ra | Chạy trong môi trường không có giao diện, hoặc thiếu plt.show() | Thêm plt.savefig("thu.png") để ghi ra tệp ảnh thay vì hiện cửa sổ |
| Phổ có nhiều vạch thay vì một | Quên nhân cửa sổ Hanning, hoặc biên độ vượt quá 1,0 gây méo | Kiểm lại dòng nhân cửa sổ; in np.max(np.abs(x)) xem có vượt 1 không |
| Tệp WAV nghe bị rè | Biên độ vượt 1,0 nên bị cắt ngọn khi ghi | Giảm hệ số 0,5 xuống, hoặc chuẩn hoá trước khi ghi |

Ba lỗi đầu là lỗi môi trường, gặp một lần rồi thôi. Lỗi cuối là lỗi *tín hiệu*, và nó sẽ còn gặp lại nhiều lần — đó là hiện tượng cắt ngọn, sẽ được bàn kỹ ở chương 5.

## 1.7 Bài tập tính toán

Sáu bài dưới đây đều giải được bằng các công thức của chương. Yêu cầu trình bày phép tính, không chỉ ghi kết quả.

***Bài 1 — bước sóng và kích thước vật cản***

Một phòng thu có cột bê tông đường kính 40 cm.

- a) Tính bước sóng của âm 100 Hz, 1.000 Hz và 8.000 Hz ở 20 °C.
- b) Với mỗi tần số, cho biết cái cột là vật cản đáng kể hay sóng đi vòng qua như không có gì.
- *Đáp số: a) 3,43 m · 0,343 m · 0,043 m. b) 100 Hz có bước sóng gấp 9 lần đường kính cột nên đi vòng qua dễ dàng; 8.000 Hz có bước sóng chỉ bằng 0,11 lần đường kính nên cột thành vật cản thật, tạo vùng tối âm phía sau.*

***Bài 2 — quy tắc 6 dB ở hiện trường***

Một micro đặt cách người nói 2 m đo được 70 dB SPL. Nhiễu nền của phòng là 45 dB SPL và không đổi theo khoảng cách.

- a) Tính tỉ số tín hiệu trên nhiễu ở vị trí hiện tại.
- b) Đưa micro về còn 0,5 m thì mức tín hiệu là bao nhiêu?
- c) Tỉ số tín hiệu trên nhiễu lúc đó cải thiện được bao nhiêu dB?
- *Đáp số: a) 70 − 45 = 25 dB. b) 82 dB SPL. c) tỉ số mới là 37 dB, cải thiện 12 dB — đúng bằng phần tín hiệu tăng lên, vì nhiễu nền không đổi.*

***Bài 3 — hai thang decibel***

Một tín hiệu có biên độ đỉnh 0,25 trong thang ±1,0.

- a) Mức đỉnh của nó là bao nhiêu dBFS?
- b) Nếu nhân dãy mẫu với 4 thì mức mới là bao nhiêu?
- c) Công suất của tín hiệu tăng bao nhiêu dB khi nhân với 4?
- *Đáp số: a) −12,04 dBFS. b) 0 dBFS — vừa đúng toàn thang. c) biên độ gấp 4 thì công suất gấp 16, tức +12,04 dB — cùng một con số với câu b, chỉ khác cách nói.*

***Bài 4 — đỉnh, hiệu dụng và hệ số đỉnh***

Cho ba tín hiệu cùng biên độ đỉnh 1,0: một sóng sin, một sóng vuông và một đoạn nhiễu trắng.

- a) Tính giá trị hiệu dụng của sóng sin và sóng vuông bằng công thức.
- b) Tính hệ số đỉnh của cả ba.
- c) Nếu chuẩn hoá cả ba về cùng mức *đỉnh* rồi nghe, cái nào nghe to nhất? Vì sao?
- *Đáp số: a) sin: $1/\sqrt{2} = 0{,}7071$; vuông: 1,0000. b) 3,01 dB · 0,00 dB · 13,1 dB. c) sóng vuông, vì hiệu dụng lớn nhất — mà độ to cảm nhận bám theo hiệu dụng chứ không bám theo đỉnh.*

***Bài 5 — tốc độ âm và độ trễ***

Một sân khấu ngoài trời có cụm loa phụ đặt cách sân khấu chính 34 m. Buổi diễn ban ngày nhiệt độ 34 °C, buổi tối còn 24 °C.

- a) Tính độ trễ cần đặt cho cụm loa phụ ở mỗi nhiệt độ.
- b) Chênh lệch giữa hai lần là bao nhiêu mili giây?

- c) Nếu cài một lần theo nhiệt độ ban ngày rồi để nguyên tới tối thì sai số ấy có đáng kể không?
- *Đáp số: a) 96,8 ms và 98,4 ms. b) 1,61 ms. c) dưới một mili giây trên quãng 34 m — nhỏ hơn ngưỡng nghe thấy, nên cài một lần là đủ. Nhưng với quãng vài trăm mét thì phải tính lại.*

***Bài 6 — phổ hài***

Một âm có tần số cơ bản 150 Hz.

- a) Liệt kê tần số của năm hài đầu tiên.
- b) Nếu đó là sóng vuông thì hài nào có mặt, hài nào không?
- c) Hài nào rơi vào vùng tai nhạy nhất (2 – 5 kHz)?
- *Đáp số: a) 150 · 300 · 450 · 600 · 750 Hz. b) chỉ các hài lẻ: 150, 450, 750 Hz… các hài chẵn bằng không. c) hài bậc 14 tới bậc 33, tức từ 2.100 tới 4.950 Hz — nghĩa là vùng tai nhạy nhất nằm ở các hài bậc cao chứ không phải tần số cơ bản.*

***Cách trình bày bài làm***

Sáu bài trên đều có đáp số in sẵn, nên điều quan trọng không phải con số cuối cùng mà là đường đi tới nó. Bài làm đạt yêu cầu cần có đủ ba thứ: công thức dùng, các con số thay vào công thức ấy, và kết quả kèm đơn vị. Một dòng chỉ ghi mỗi kết quả, dù đúng, cũng không cho thấy người làm hiểu vì sao.

**Ba lỗi hay gặp khi làm bài tập nhóm này**

*Lẫn đơn vị*. Bước sóng ra mét, tần số ra héc, thời gian ra giây hoặc mili giây. Sai đơn vị thì con số lệch nghìn lần mà nhìn vẫn thấy "hợp lý". Cách kiểm nhanh: viết đơn vị vào ngay cạnh mỗi số khi thay vào công thức, rồi rút gọn xem có ra đúng đơn vị cần tìm.

*Dùng nhầm hệ số decibel*. Hệ số 20 cho biên độ và điện áp, hệ số 10 cho công suất và năng lượng. Trong cả chương này và phần lớn cuốn sách, cái đang xử lý là dãy mẫu, tức biên độ, nên hệ số là 20.

*Cộng decibel như cộng số thường*. Decibel là thang lô-ga-rít: hai nguồn 60 dB cùng phát không cho 120 dB mà cho khoảng 63 dB. Khi thấy một phép cộng decibel trong bài làm, hãy dừng lại kiểm tra xem nó có thật sự đúng không.

Người muốn tự kiểm tra có thể mở tệp sigtools.py đi kèm sách: mọi đáp số ở trên đều do các hàm trong tệp ấy sinh ra, và thay tham số của riêng mình vào là ra con số cho trường hợp của mình. Nhưng nên tính tay trước rồi mới đối chiếu, chứ không nên làm ngược lại.

## 1.8 Câu hỏi ôn tập và bài thảo luận

### 1.8.1 Câu hỏi ôn tập

- 1. Nêu ba yếu tố phải có đủ để có âm thanh. Thiếu yếu tố nào thì hoàn toàn không có âm?

- 2. Vì sao âm thanh không truyền được trong chân không, còn ánh sáng thì được?
- 3. Phân biệt sóng dọc và sóng ngang. Âm thanh trong không khí thuộc loại nào?
- 4. Cái gì thật sự lan đi trong không gian khi có sóng âm — các phần tử không khí hay trạng thái nén?
- 5. Viết công thức của một sóng sin và cho biết ý nghĩa từng tham số.
- 6. Tốc độ âm trong không khí phụ thuộc yếu tố nào? Ở 20 °C là bao nhiêu?
- 7. Vì sao âm truyền trong thép nhanh hơn trong không khí, dù thép đặc hơn nhiều?
- 8. Viết công thức bước sóng. Bước sóng của 20 Hz và 20 kHz chênh nhau bao nhiêu lần?
- 9. Vì sao tấm mút tiêu âm mỏng không hút được âm trầm?
- 10. Phân biệt tần số và cao độ. Quan hệ giữa chúng là quan hệ cộng hay nhân?
- 11. Phân biệt biên độ đỉnh, giá trị hiệu dụng và hệ số đỉnh. Đại lượng nào gần với độ to cảm nhận nhất?
- 12. Vì sao decibel có hai công thức? Khi nào dùng hệ số 20, khi nào dùng hệ số 10?
- 13. Gấp đôi biên độ là bao nhiêu dB? Gấp đôi công suất là bao nhiêu dB?
- 14. Phát biểu quy tắc 6 dB theo khoảng cách và nêu một cách dùng nó ở hiện trường.
- 15. Pha là gì? Vì sao nghe một sóng sin đơn độc thì không nhận ra pha, nhưng khi cộng hai sóng thì pha quyết định tất cả?
- 16. Mô tả hiện tượng đảo pha khi thu bằng hai micro. Triệu chứng nhận biết là gì?
- 17. Vì sao cùng một nốt mà piano và sáo nghe khác nhau? Trả lời bằng khái niệm phổ hài.
- 18. Sóng vuông có những hài nào? Sóng răng cưa thì sao?
- 19. Vì sao gõ trống không cho cảm giác cao độ rõ ràng như gảy đàn?
- 20. Dải nghe được của tai người là bao nhiêu? Nó thay đổi thế nào theo tuổi?
- 21. Vì sao điện thoại đời cũ chỉ truyền 300 – 3.400 Hz mà vẫn nghe rõ lời?
- 22. Tai nhạy nhất ở vùng tần số nào? Nêu một hệ quả thực hành của điều đó.
- 23. Mô tả hiện tượng che lấp theo tần số. Vì sao búp che lấp không đối xứng?
- 24. Giải thích bằng lời của chính mình: vì sao MP3 bỏ được phần lớn dữ liệu mà tai vẫn thấy như cũ?
- 25. Đọc một phổ đồ: trục ngang, trục dọc và độ sáng lần lượt biểu diễn đại lượng gì?

### 1.8.2 Bài thảo luận

Hai tình huống dưới đây không có đáp án đúng duy nhất. Yêu cầu lập luận bằng các khái niệm đã học trong chương.

### *Tình huống 1 — bản thu bị mỏng khi trộn hai micro*

Một nhóm thu tiếng đàn ghi-ta bằng hai micro: một đặt trước lỗ thoát âm, một đặt gần cần đàn. Nghe riêng từng kênh thì cả hai đều dày và đầy đặn. Nhưng trộn hai kênh lại thành mono thì tiếng mỏng hẳn đi, mất phần trầm.

- Nêu nguyên nhân có khả năng nhất và giải thích bằng khái niệm đã học.
- Vì sao chỉ phần *trầm* bị mất mà phần cao thì không? Gợi ý: nghĩ về bước sóng so với khoảng cách giữa hai micro.
- Nêu hai cách khắc phục, một cách ở khâu đặt micro và một cách ở khâu xử lý.

### *Tình huống 2 — chọn tần số lấy mẫu cho một dự án*

Một nhóm làm âm thanh cho trò chơi trên điện thoại. Trưởng nhóm đề nghị thu và lưu toàn bộ ở 96 kHz, 24 bit, lập luận rằng "chất lượng cao nhất thì không bao giờ thiệt".

- Dựa vào dải nghe được đã học, phần nào của lập luận ấy đúng và phần nào đáng bàn lại?
- Việc lưu ở 96 kHz có làm âm thanh nghe hay hơn trên loa điện thoại không? Vì sao?
- Nêu một lý do *kỹ thuật* chính đáng để vẫn thu ở tần số cao, dù sản phẩm cuối phát ở tần số thấp hơn. Gợi ý: nghĩ về các bước xử lý nằm giữa lúc thu và lúc phát.

Câu hỏi thứ ba sẽ được trả lời đầy đủ ở chương 3. Ở đây chỉ cần đưa ra phỏng đoán có lập luận.

### *Cách lập luận trong hai tình huống trên*

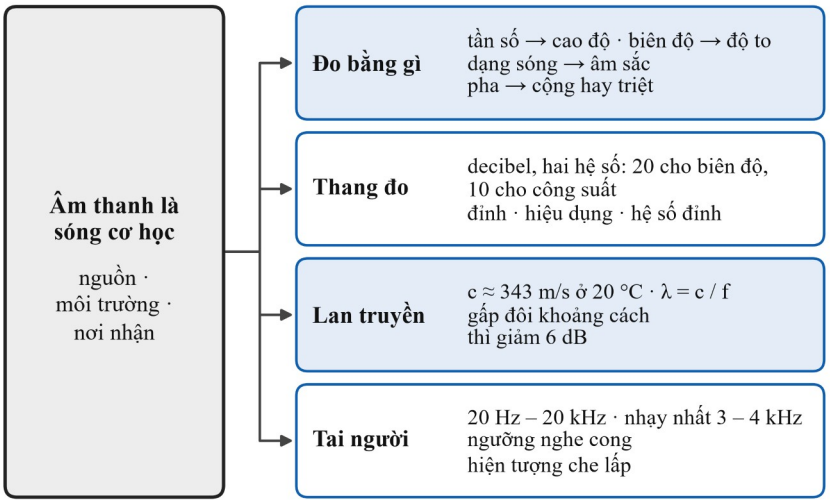
Hai tình huống này không có đáp án dạng một con số, nên điều cần chú trọng là chất lượng lập luận. Một lập luận tốt trong nghề âm thanh thường đi theo bốn bước, và người đọc nên tập viết theo thứ tự ấy.

- **Mô tả hiện tượng bằng đại lượng, không bằng cảm giác.** "Nghe mỏng" là cảm giác; "mất năng lượng ở vùng vài trăm héc" là đại lượng. Chỉ khi đã dịch sang đại lượng thì mới kiểm tra được.
- **Nêu cơ chế vật lý có thể gây ra nó.** Ở tình huống 1, cơ chế là hai bản sao của cùng một nguồn lệch pha nhau vì đường đi tới hai micro dài ngắn khác nhau.
- **Đưa ra một phép thử phân biệt được.** Nếu cơ chế đúng thì tắt một micro đi, tiếng phải hết mỏng. Một giả thuyết không kèm phép thử thì chưa dùng được.
- **Đề xuất cách xử lý, kèm cái giá phải trả.** Mọi cách khắc phục đều có giá: dời micro thì đổi âm sắc, đảo pha thì sai nếu hai nguồn không thật sự đối pha. Nêu được cái giá cho thấy đã hiểu, chứ không phải chỉ đọc một quy tắc ghi nhớ.

Bốn bước này sẽ được dùng lại ở mọi chương thực hành sau, và ở phần gỡ lỗi âm thanh của các dự án thực hành. Nên tập ngay từ chương đầu.

## 1.9 Tóm tắt và chuẩn bị cho chương sau

### 1.9.1 Bản đồ khái niệm




*Bốn nhánh này quay lại ở mọi chương sau. Nhánh thang đo dùng nhiều nhất — gần như chương nào cũng gặp.*

*Hình 1.29 Bản đồ khái niệm chương 1 — bốn nhánh quay lại ở mọi chương sau*

**Năm điều phải nhớ sau chương này**

Âm thanh là *sóng cơ học* — cần nguồn, môi trường và nơi nhận. Cái lan đi là trạng thái nén, không phải vật chất.

Bốn đại lượng gốc: tần số quyết định cao độ, biên độ quyết định độ to, dạng sóng quyết định âm sắc, pha quyết định chuyện cộng hay triệt.

Decibel có *hai* hệ số: 20 cho biên độ, 10 cho công suất. Trong cuốn sách này gần như luôn dùng hệ số 20.

Âm sắc = tỉ lệ các *hài*. Cùng tần số cơ bản mà khác tỉ lệ hài thì khác nhạc cụ.

Tai có hai *điểm mù*: ngưỡng nghe cong theo tần số, và hiện tượng che lấp. Toàn bộ chuyện nén có tổn hao dựa trên hai điểm mù ấy.

### 1.9.2 Bảng thuật ngữ Anh – Việt

*Bảng 1.11 Thuật ngữ chương 1*

| Tiếng Anh | Tiếng Việt | Ghi chú |
|---|---|---|
| Longitudinal wave | Sóng dọc | phần tử dao động theo hướng sóng đi |
| Compression / Rarefaction | Nén / Giãn | hai pha của sóng âm trong không khí |
| Amplitude | Biên độ | độ lệch lớn nhất khỏi mức yên tĩnh |
| Frequency | Tần số | số chu kỳ trong một giây, đơn vị Hz |
| Wavelength | Bước sóng | $\lambda = c / f$ |

| Tiếng Anh | Tiếng Việt | Ghi chú |
|---|---|---|
| Phase | Pha | vị trí trong chu kỳ, đo bằng độ hoặc radian |
| Timbre | Âm sắc | do tỉ lệ các hài quyết định |
| Harmonic | Hài | thành phần ở bội số nguyên của tần số cơ bản |
| Fundamental frequency | Tần số cơ bản | ký hiệu $f_0$ |
| RMS (root mean square) | Giá trị hiệu dụng | gần với độ to cảm nhận hơn đỉnh |
| Crest factor | Hệ số đỉnh | tỉ số đỉnh trên hiệu dụng, tính bằng dB |
| dBFS (decibels full scale) | Decibel toàn thang | thang trong máy tính, 0 dBFS là biên độ 1,0 |
| SPL (sound pressure level) | Mức áp suất âm | thang trong không khí, đơn vị dB SPL |
| Absolute threshold of hearing | Ngưỡng nghe tuyệt đối | mức tối thiểu để nghe thấy, phụ thuộc tần số |
| Masking | Che lấp | âm mạnh nâng ngưỡng nghe của vùng lân cận |
| Spectrum | Phổ | phân bố năng lượng theo tần số |
| Spectrogram | Phổ đồ | phổ thay đổi theo thời gian |
| Đường bao biên độ | Amplitude envelope | Cách biên độ của một nốt thay đổi từ lúc bắt đầu tới lúc tắt. |
| Khởi công / Suy giảm / Duy trì / Tắt dần | Attack / Decay / Sustain / Release (ADSR) | Bốn giai đoạn của mô hình đường bao ADSR. |
| Đường đẳng âm | Equal-loudness contour | Đường nối các âm nghe to bằng nhau ở các tần số khác nhau. |
| Phon | Phon | Đơn vị độ to: N phon là to bằng âm 1 kHz ở N dB SPL. |
| Sone | Sone | Đơn vị độ to theo tỉ lệ: 1 sone ứng với 40 phon, gấp đôi độ to thì gấp đôi số sone. |

### 1.9.3 Chuẩn bị cho chương 2

Chương 2 nói về âm thanh *analog* và chuỗi thiết bị: micro biến dao động cơ thành điện thế nào, tín hiệu đi qua những khối nào trước khi thành tệp, và vì sao người ta phải chuyển sang số.

- **Việc phải làm:** hoàn tất cài môi trường và chạy được đoạn mã ở mục 1.6.2. Chương 2 sẽ bắt đầu bằng việc đọc ghi tệp WAV, nên môi trường phải chạy được trước.
- **Bài tập:** chọn ba âm thanh trong môi trường xung quanh — một trầm, một trung, một cao. Với mỗi âm, viết một đoạn ngắn mô tả theo bốn đại lượng đã học: ước lượng tần số, so sánh biên độ, nhận xét âm sắc.
- **Nên làm thêm:** mở Audacity, thu giọng nói của chính mình trong một câu, rồi xem cả dạng sóng lẫn phổ đồ. Thử tìm xem chỗ nào là nguyên âm, chỗ nào là phụ âm — sẽ thấy chúng trông rất khác nhau trên phổ đồ.

### 1.9.4 Bảng tự kiểm trước khi sang chương 2

Chương 2 xây trực tiếp trên chương này, nên trước khi sang, hãy tự trả lời tám câu dưới đây mà không mở lại tài liệu. Câu nào ngập ngừng thì quay lại đúng mục ghi trong ngoặc.

- Kể ba yếu tố phải có đủ để có âm thanh, và nói vì sao thiếu môi trường thì không còn âm thanh (1.2.1).
- Cho nhiệt độ phòng, tính được tốc độ âm và bước sóng của một tần số cho trước (1.2.3, 1.2.4).
- Gọi tên bốn đại lượng gốc và nói mỗi đại lượng quyết định cảm giác nghe nào (1.3).
- Nói được khi nào dùng hệ số 20 và khi nào dùng hệ số 10 trong công thức decibel (1.3.3).
- Nhớ được quy tắc 6 dB và áp dụng nó cho một khoảng cách bất kỳ (1.3.4).
- Giải thích được vì sao cùng một nốt mà piano và sáo nghe khác nhau, bằng khái niệm phổ hài (1.4.3).
- Nói được hai điểm mù của tai người và liên hệ chúng với nén có tổn hao (1.5.4).
- Chạy được đoạn mã ở mục 1.6.2 trên máy của mình và đọc được ba con số nó in ra (1.6.2).

> **Nếu còn đúng một việc chưa xong**
> Thì việc ấy nên là cài xong môi trường. Kiến thức thiếu chỗ nào có thể vá dần trong chương 2, nhưng máy chưa chạy được Python và soundfile thì cả chương 2 sẽ chỉ ngồi nhìn, vì chương ấy bắt đầu ngay bằng việc đọc và ghi tệp WAV.

## Tài liệu tham khảo của chương

### *Sách và giáo trình*


- Fletcher, H., Munson, W. A. — Loudness, its definition, measurement and calculation, Journal of the Acoustical Society of America, 1933. Bài gốc của đường đẳng âm.
- Stevens, S. S. — A scale for the measurement of a psychological magnitude: loudness, Psychological Review, 1936. Nguồn của thang sone.

- Rossing, T. D., Moore, F. R., Wheeler, P. A. — The Science of Sound. Sách nền về vật lý âm thanh và thính giác; nguồn của phần sóng cơ học và tốc độ truyền âm.
- Everest, F. A., Pohlmann, K. C. — Master Handbook of Acoustics. Nguồn của phần bước sóng, suy giảm theo khoảng cách và âm học phòng.
- Zwicker, E., Fastl, H. — Psychoacoustics: Facts and Models. Nguồn của phần ngưỡng nghe và hiện tượng che lấp.
- Smith, J. O. — Mathematics of the Discrete Fourier Transform. Nguồn của phần phổ và phân tích hài; sẽ dùng nhiều hơn ở chương 4.

***Tiêu chuẩn***

- ISO 226:2003 — Acoustics: Normal equal-loudness-level contours. Nguồn chuẩn của đường đẳng âm và ngưỡng nghe; số liệu ở mục 1.5.2 lấy từ bảng của chuẩn này. ISO 532 — Acoustics: Methods for calculating loudness, tính độ to sone cho âm phức tạp.
- IEC 61672 — Electroacoustics: Sound level meters. Quy định cách đo mức áp suất âm.

***Một lưu ý về các con số trong bài***

> **Cách đọc các con số**
>
> Các con số có nguồn tiêu chuẩn — dải nghe được, đường ngưỡng nghe, tốc độ âm — được nêu đúng như tài liệu gốc quy định.
>
> Các con số mô tả nhạc cụ và nguồn âm là giá trị *tiêu biểu* dùng để so sánh, không phải số đo của một nhạc cụ cụ thể.
>
> Mọi con số tính toán trong bài và mọi đáp số bài tập đều sinh ra từ mô-đun sigtools.py, đã qua 54 phép tự kiểm. Người đọc thay tham số của riêng mình vào là ra con số cho trường hợp của mình.