MỤC LỤC

Chương 1. Bản chất vật lý của âm thanh và cơ chế cảm nhận thính giác.....7
1.1 Bản chất vật lý của âm thanh và cơ chế lan truyền sóng cơ học.....8
1.2 Bốn đại lượng đặc trưng của sóng âm.....13
1.3 Âm học phòng.....24
1.4 Đơn vị decibel và các quy chuẩn đo lường.....32
1.5 Cơ chế phát âm giọng nói và các dải tần cộng hưởng (Formants).....42
1.6 Sinh lý thính giác và tâm lý âm học.....44
1.7 Mô phỏng kỹ thuật và cầu nối lập trình âm thanh số.....51
1.8 Hệ thống câu hỏi và bài tập.....55
1.9 Bảng đối chiếu thuật ngữ chuyên ngành Anh – Việt.....62

DANH MỤC HÌNH VẼ

Hình 1.1 Bản chất sóng dọc: miền nén và miền giãn của các phần tử không khí, đối chiếu với đồ thị biến thiên áp suất âm thanh.....9
Hình 1.2 Biên độ đỉnh, đỉnh – đỉnh, giá trị hiệu dụng RMS và ngưỡng méo cắt đỉnh của hệ thống.....15
Hình 1.3 Dạng sóng âm sin 440 Hz phóng to đến mức thấy từng chu kỳ, kèm thước thời gian (Audacity).....15
Hình 1.4 Sáu phân vùng tần số của dải nghe được và các mốc thường dùng trong xử lý âm thanh.....17
Hình 1.5 Phổ vạch hài của một âm phức hợp (cơ bản 220 Hz) trong hộp thoại Frequency Analysis của Audacity.....18
Hình 1.6 Âm sắc: phổ hài Fourier của nốt La 440 Hz ở ba nguồn âm khác nhau và mô hình đường bao biên độ ADSR.....19
Hình 1.7 Tương tác pha giữa hai sóng âm và đáp tuyến lọc lược (Comb Filtering) khi trộn tín hiệu với bản sao trễ 1 ms.....23
Hình 1.8 Phòng đo vô hưởng với các khối nêm hấp thụ âm phủ tường và micro đo trên giá đỡ.....25
Hình 1.9 Bốn kiểu tương tác của sóng âm với bề mặt: phản xạ gương, khuếch tán, hội tụ và hấp thụ.....26
Hình 1.10 Sóng đứng trong phòng kín: vị trí nút và bụng sóng, phổ các mode và giải pháp bố trí bẫy âm trầm.....28
Hình 1.11 Phòng điều khiển phòng thu có tấm tán âm dạng thanh và vách vát góc.....30
Hình 1.12 Hệ đường cong tiêu chí tiếng ồn nền NC và mức yêu cầu đối với từng loại không gian sản xuất.....31
Hình 1.13 Máy đo mức âm chính xác với micro đo ở đầu máy và màn hình hiển thị dB(A).....34
Hình 1.14 Thang mức áp suất âm dB SPL từ ngưỡng nghe tuyệt đối đến ngưỡng đau, kèm giới hạn tiếp xúc nghề nghiệp.....35
Hình 1.15 Đồng hồ đo độ to LUFS (Loudness Meter) trong Adobe Audition; nguồn âm là tệp thoại mẫu tổng hợp.....38
Hình 1.16 Định luật bình phương nghịch đảo: mức âm giảm 6 dB mỗi lần nhân đôi cự ly, và hệ quả với cự ly đặt micro.....41
Hình 1.17 Thang decibel toàn dải dBFS, khoảng đệm an toàn và các mốc chuẩn trong quy trình gain staging.....42
Hình 1.18 Cơ quan phát âm của người theo mô hình nguồn – bộ lọc: dây thanh quyết định cao độ, khoang cộng hưởng quyết định nguyên âm.....43
Hình 1.19 Phổ đồ của đoạn thoại mẫu trong Adobe Audition; nguồn âm là tệp thoại tổng hợp, trục đứng là tần số.....43
Hình 1.20 Cấu tạo ba khoang tai, đường đi của tín hiệu và cơ chế phối hợp trở kháng âm học.....46
Hình 1.21 Hệ đường cong đẳng âm dựng theo bảng tham số chuẩn ISO 226 và vùng nghe của tai người.....47
Hình 1.22 Đặc tuyến mạch lọc trọng số dBA (cắt trầm mạnh) đối chiếu dBC (gần như phẳng).....48
Hình 1.23 Đường cong ngưỡng che lấp trong miền tần số — cơ sở vật lý của mọi thuật toán nén có tổn hao.....49
Hình 1.24 Đường cong hiệu ứng ưu tiên (Haas) và sơ đồ cài đặt độ trễ cho hệ loa vòm 50

DANH MỤC BẢNG BIỂU

Bảng 1.1 Vận tốc truyền âm trong các môi trường vật chất.....10
Bảng 1.2 Bước sóng vật lý của âm thanh và Hệ quả kỹ thuật trong thiết kế âm học.....12
Bảng 1.3 Sáu phân vùng tần số thường dùng trong xử lý âm thanh.....16
Bảng 1.4 Hệ số hấp thụ âm thanh α của một số vật liệu thường gặp.....25
Bảng 1.5 Thời gian âm vang RT60 khuyến nghị cho một số loại không gian.....29
Bảng 1.6 Danh mục các mốc quy chiếu Decibel chuẩn mực trong kỹ thuật Audio.....33
Bảng 1.7 Ba đại lượng công suất, cường độ, thanh áp và công thức mức tương ứng....35
Bảng 1.8 Tiêu chuẩn đo độ to LUFS và Trần True Peak trên các nền tảng phát hành. .39
Bảng 1.9 Bảng tra nhanh mức suy hao âm thanh theo khoảng cách.....40
Bảng 1.10 Phân loại các chất giọng và đặc tính xử lý âm thanh.....44
Bảng 1.11 Các ngưỡng chênh lệch cảm nhận của thính giác và Ý nghĩa nghiệp vụ.....48
Bảng 1.12 Tiêu chuẩn giới hạn thời gian tiếp xúc tiếng ồn an toàn nghề nghiệp (NIOSH).....50
Bảng 1.13 Thuật ngữ chuyên ngành âm thanh Anh – Việt.....62

DANH MỤC KÝ HIỆU VÀ CHỮ VIẾT TẮT

Ký hiệu viết tắt	Thuật ngữ tiếng Anh đầy đủ	Ý nghĩa chuyên môn trong kỹ thuật Audio - Video
ADSR	Attack - Decay - Sustain - Release	Mô hình đường bao biên độ 4 giai đoạn theo thời gian trong thiết kế tiếng động
AES	Audio Engineering Society	Hội Kỹ sư Âm thanh Quốc tế — tổ chức ban hành các tiêu chuẩn âm thanh chuyên nghiệp
DAW	Digital Audio Workstation	Trạm làm việc âm thanh số (Pro Tools, DaVinci Resolve Fairlight, Nuendo, Reaper)
dB	Decibel	Đơn vị logarit biểu diễn tỷ số giữa hai đại lượng biên độ hoặc công suất
dB SPL	Decibel Sound Pressure Level	Mức áp suất âm thanh trong không khí quy chiếu theo ngưỡng nghe 20 µPa
dBFS	Decibels Relative to Full Scale	Thang đo decibel số toàn dải; 0 dBFS là trần tuyệt đối của hệ thống kỹ thuật số
dBTP	Decibels True Peak	Mức đỉnh thực của tín hiệu số có tính toán hiện tượng liên mẫu (inter-sample peaks)
dBA / dBC	A-weighted / C-weighted Decibels	Mức đo decibel qua mạch lọc bù trừ đặc tuyến độ nhạy của tai người
EBU	European Broadcasting Union	Hiệp hội Phát thanh Truyền hình Châu Âu (tác giả tiêu chuẩn EBU R128)
HRTF	Head-Related Transfer Function	Hàm truyền liên quan đến đầu người — cơ sở định vị không gian 3 chiều và âm thanh nhị nhĩ
ILD	Interaural Level Difference	Độ chênh lệch mức áp suất âm thanh giữa hai tai (chủ đạo ở tần số cao > 1,5 kHz)
ITD	Interaural Time Difference	Độ chênh lệch thời gian âm thanh truyền đến hai tai (chủ đạo ở tần số thấp < 1,5 kHz)
ITU-R	ITU Radiocommunication Sector	Phân ban Vô tuyến của Liên minh Viễn thông Quốc tế (tác giả chuẩn BS.1770)
LKFS / LUFS	Loudness K-weighted / Units Full Scale	Đơn vị đo lường độ to cảm nhận chuẩn hóa theo ITU-R BS.1770
LRA	Loudness Range	Dải độ to — chỉ số mô tả độ động giữa đoạn to và đoạn nhỏ của chương trình (đơn vị LU)

Ký hiệu viết tắt	Thuật ngữ tiếng Anh đầy đủ	Ý nghĩa chuyên môn trong kỹ thuật Audio - Video
NC / NR	Noise Criteria / Noise Rating	Đường cong tiêu chí đánh giá mức độ ồn nền của phòng thu và trường quay
NIOSH	National Institute for Occupational Safety and Health	Viện Quốc gia về An toàn và Sức khỏe Nghề nghiệp Hoa Kỳ
RMS	Root Mean Square	Giá trị căn bậc hai trung bình bình phương (năng lượng hiệu dụng của tín hiệu)
RT60	Reverberation Time 60 dB	Thời gian để năng lượng âm vang suy giảm 60 dB sau khi ngắt nguồn âm
SMPTE	Society of Motion Picture and Television Engineers	Hiệp hội Kỹ sư Điện ảnh và Truyền hình Quốc tế
SNR	Signal-to-Noise Ratio	Tỷ số tín hiệu trên tạp âm (đo bằng dB)

CHƯƠNG 1
BẢN CHẤT VẬT LÝ CỦA ÂM THANH VÀ CƠ CHẾ CẢM NHẬN THÍNH GIÁC

Chương 1 xây dựng nền tảng vật lý và sinh lý học của toàn bộ cuốn sách. Mọi thiết bị, phần mềm và quy trình được trình bày ở các chương sau, từ microphone, loa, bộ chuyển đổi tương tự - số A/D cho đến các công cụ trí tuệ nhân tạo (AI) hậu kỳ, đều chỉ là những phương thức khác nhau để thao tác trên các đại lượng được định nghĩa ở đây. Người đọc nắm vững chương này có thể chẩn đoán chính xác nguyên nhân của phần lớn sự cố âm thanh khi thu âm và hòa âm, thay vì phải phỏng đoán.

Mục tiêu của chương

Sau khi đọc xong chương này, người đọc có thể:

1. **Về kiến thức khoa học (Knowledge):** Trình bày chính xác bản chất sóng cơ học và sóng dọc của âm thanh trong chất khí; phân tích ba điều kiện tiên quyết cho sự hình thành âm thanh; định nghĩa và giải thích ý nghĩa vật lý của bốn đại lượng cơ bản: biên độ (Peak, RMS, Crest Factor), tần số (quãng tám, decade), âm sắc (chuỗi hài Fourier, mô hình đường bao ADSR) và góc pha; nắm vững các hiện tượng âm học phòng (dội âm, âm vang, sóng đứng, tiếng vang); hiểu bản chất logarit của thang đo Decibel (dB SPL, dBu, dBFS, dBTP, dBA, dBC), định luật suy hao khoảng cách, cơ chế phát âm formant và các quy luật tâm lý âm học hiện đại (ISO 226:2023, che lấp âm, hiệu ứng Haas).
2. **Về kỹ năng chuyên môn nghề nghiệp (Skills):** Vận dụng công thức vật lý để tính chính xác bước sóng và vận tốc âm theo nhiệt độ môi trường; áp dụng định luật suy hao khoảng cách (-6 dB khi nhân đôi cự ly) để chọn khoảng cách đặt micro và thiết lập Gain tiền khuếch đại tối ưu nhằm bảo toàn SNR; tính tần số sóng đứng (Room Modes) để xử lý tiêu âm góc phòng dựng; áp dụng quy tắc 3:1 để tránh triệt tiêu pha khi thu nhiều nguồn âm bằng nhiều micro; tính độ trễ (delay) cho hệ thống loa theo hiệu ứng Haas; đo lường và hiệu chỉnh độ to theo tiêu chuẩn ITU-R BS.1770 (LUFS/LKFS) và kiểm soát True Peak chống méo trên các nền tảng phát hành số (Netflix, YouTube, Spotify, EBU R128).
3. **Về thói quen làm việc: Chẩn đoán nguyên nhân gốc của sự cố âm thanh dựa trên định luật vật lý thay vì phỏng đoán cảm tính; tuân thủ các quy tắc bảo vệ thính giác** theo khuyến nghị NIOSH và mức kiểm chuẩn âm lượng phòng hòa âm (85 dBC SPL).

1.1 Bản chất vật lý của âm thanh và cơ chế lan truyền sóng cơ học

1.1.1 Khái niệm và ba điều kiện tồn tại tiên quyết của âm thanh

Dưới góc độ vật lý học thuần túy, âm thanh là sự lan truyền của dao động cơ học dưới dạng sóng áp suất qua một môi trường vật chất có tính đàn hồi. Dưới góc độ sinh lý học, âm thanh là cảm giác thính giác được não bộ con người tái tạo và cảm thụ khi màng nhĩ rung lên trước tác động của những biến thiên áp suất đó.

Sự phân biệt giữa hai định nghĩa này mang tính bản lề cho kỹ thuật audio: thiết bị đo đạc cung cấp các đại lượng vật lý khách quan (như Pascal, Watt), trong khi người nghe tiếp nhận bằng cảm nhận chủ quan (độ to, cao độ, âm sắc). Để hiện tượng âm thanh hình thành, lan truyền và được ghi nhận, cần hội tụ đầy đủ ba yếu tố tiên quyết:

1. **Nguồn tạo ra âm thanh (Acoustic source):** Là một hệ vật thể cơ học thực hiện dao động quanh vị trí cân bằng bền dưới tác động của ngoại lực. Âm thanh trong thực tế là sản phẩm của các hiện tượng va chạm, cọ xát cơ học giữa các vật thể rắn, lỏng hoặc khí: dây thanh quản của người nói rung lên dưới áp lực luồng khí phổi, luồng khí phụt qua miệng sáo, tiếng bước chân nện xuống sàn gỗ, hay màng loa nén dao động.
2. **Môi trường truyền dẫn đàn hồi (Transmission medium):** Rung động từ nguồn âm phải được truyền qua môi trường vật chất liên tục có tính đàn hồi (chất khí, chất lỏng hoặc chất rắn) để tạo ra biến thiên áp suất. Sóng âm không thể truyền qua chân không vì thiếu các hạt phân tử làm trung gian truyền động năng.
3. **Nơi tiếp nhận dao động âm thanh (Receiver):** Là cơ quan hoặc cảm biến có khả năng tiếp nhận và đáp ứng trung thực với sự biến thiên áp suất của môi trường. Trong đời sống, đó là màng nhĩ và cơ quan thính giác của con người; trong kỹ thuật thu âm, đó là màng rung của microphone hoặc đầu đo thủy âm (hydrophone) khi thu dưới nước.

1.1.2 Cơ chế sóng cơ học và bản chất sóng dọc trong chất khí

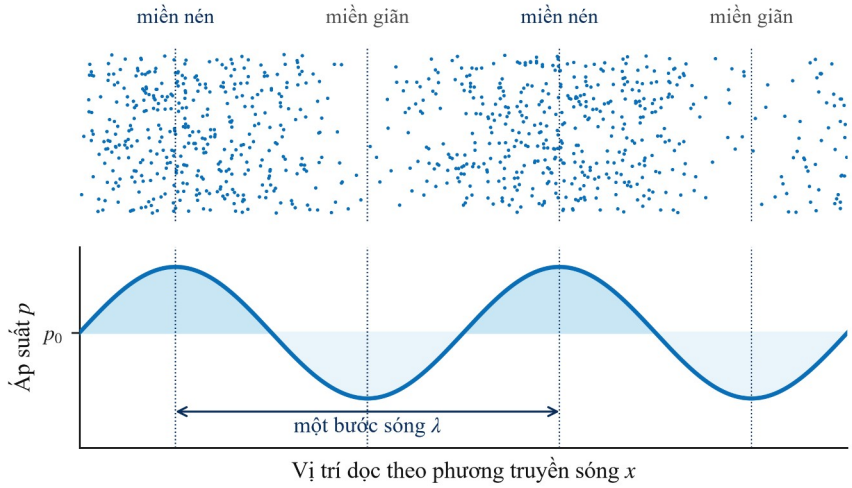
Trong môi trường chất khí và chất lỏng, sóng âm thuần túy là sóng dọc (longitudinal wave). Khi nguồn âm dao động, nó đẩy các phần tử không khí lân cận rời khỏi vị trí cân bằng ban đầu, còn lực đàn hồi của môi trường có xu hướng kéo phần tử đó trở lại vị trí cũ. Do quán tính cơ học, phần tử không khí chuyển động vượt qua vị trí cân bằng sang phía đối diện, rồi tiếp tục chịu tác dụng của lực đàn hồi kéo ngược trở lại.

Quá trình dao động tịnh tiến này được truyền lan sang các phần tử kế tiếp thông qua va chạm phân tử, tạo nên sự lan truyền xen kẽ giữa các miền không khí bị nén lại (Compression, nơi áp suất tức thời cao hơn áp suất khí quyển tĩnh P_0) và các miền không khí bị giãn ra (Rarefaction, nơi áp suất tức thời thấp hơn áp suất khí quyển tĩnh). Đường sóng hình sin quen thuộc hiển thị trên màn hình phần mềm xử lý âm thanh (DAW) không phải là hình dạng chuyển động uốn lượn của không

khí trong không gian, mà là đồ thị biểu diễn sự biến thiên độ chênh áp suất âm thanh theo trục thời gian tại một tọa độ xác định.

Một nguyên lý cơ học quan trọng thường bị ngộ nhận: sóng âm truyền năng lượng chứ không truyền vật chất. Các phần tử không khí chỉ dao động quanh vị trí cân bằng cố định với biên độ dịch chuyển rất nhỏ. Ở mức áp suất danh định 94 dB SPL (tương đương 1 Pa) tại tần số 1 kHz, biên độ dao động của phân tử không khí chỉ vào khoảng 0,07 µm, nhỏ hơn cả bước sóng ánh sáng khả kiến. Tại ngưỡng nghe của con người (0 dB SPL, tương đương 20 µPa), khoảng cách dịch chuyển này thu hẹp xuống mức 10^{-11} m, tức nhỏ hơn cả đường kính của một nguyên tử hydro. Độ nhạy này vừa là đặc điểm nổi bật của cơ quan thính giác con người, vừa là lý do màng rung của microphone phải được gia công rất mỏng và rất nhẹ (chỉ vài micromét).

Trong môi trường chất rắn (như sàn nhà, tường gạch, giá đỡ thiết bị), ngoài sóng dọc còn xuất hiện sóng ngang (transverse wave) và sóng bề mặt (Rayleigh wave), vì các phân tử chất rắn có thể dao động vuông góc với phương truyền sóng. Chính các xung chấn cơ học truyền qua sàn và giá đỡ này là nguồn gốc của các tạp âm cơ học tần số cực thấp (tiếng bước chân, tiếng xe cộ chạy ngoài đường rung móng nhà) xâm nhập vào micro thông qua thân giá đỡ. Do đó, microphone thu âm thường được gắn trên các bộ giảm chấn đàn hồi (Shockmount) để cô lập cơ học khỏi giá đỡ.



Hình 1.1 Bản chất sóng dọc: miền nén và miền giãn của các phần tử không khí, đối chiếu với đồ thị biến thiên áp suất âm thanh

1.1.3 Vận tốc truyền âm, công thức nhiệt động học và trở kháng âm học

Vận tốc truyền sóng âm c (speed of sound) là tốc độ lan truyền pha dao động qua môi trường đàn hồi. Về mặt nhiệt động học lý thuyết, sự lan truyền sóng âm trong chất khí là một quá trình đoạn nhiệt (adiabatic process) diễn ra cực nhanh, không kịp trao đổi nhiệt với môi trường xung quanh. Vận tốc âm trong khí lý tưởng được xác định chính xác bởi công thức Newton-Laplace [9]:

$$c = \sqrt{\frac{\gamma R T}{M}} \quad \text{(m/s)} \tag{1.1}$$

Trong đó: γ là chỉ số đoạn nhiệt của chất khí (với không khí khô γ ≈ 1,40); R = 8,314 J/(mol·K) là hằng số khí lý tưởng; T là nhiệt độ tuyệt đối tính bằng Kelvin (T = θ + 273,15, với θ là nhiệt độ Celsius); M là khối lượng mol trung bình của không khí (M ≈ 0,02896 kg/mol). Từ công thức nền tảng này, có thể rút ra ba kết luận thực tiễn:

1. **Tính độc lập với áp suất:** Vận tốc âm không phụ thuộc vào áp suất khí quyển ở điều kiện bình thường, vì khi áp suất khí quyển tăng hay giảm (do thay đổi thời tiết hoặc độ cao địa hình), mật độ không khí cũng biến thiên tỷ lệ thuận tương ứng và hai tác động triệt tiêu lẫn nhau trong phương trình.
2. **Tính độc lập với tần số:** Vận tốc âm không phụ thuộc vào tần số trong dải âm thanh nghe được, nên âm trầm và âm bổng lan truyền với cùng một tốc độ trong không khí. Nếu vận tốc âm phụ thuộc tần số, âm thanh của dàn nhạc giao hưởng nghe từ xa sẽ bị 'lệch nhịp' và mất cấu trúc tiết tấu.
3. **Ảnh hưởng của độ ẩm:** Độ ẩm không khí làm tăng nhẹ vận tốc âm, bởi khối lượng mol của phân tử hơi nước (H_2O ≈ 0,018 kg/mol) nhỏ hơn khối lượng mol trung bình của không khí khô; độ ẩm tương đối 100% ở nhiệt độ cao có thể làm tăng vận tốc âm khoảng +0,1% đến +0,35%, điều cần lưu ý khi căn chỉnh hệ thống âm thanh ngoài trời.

Trong thực hành, người ta thường áp dụng công thức thực nghiệm tuyến tính hóa xấp xỉ, thuận tiện và có độ chính xác cao trong dải nhiệt độ từ -20°C đến +50°C:

$$c \approx 331{,}3 + 0{,}606\,\theta \quad \text{(m/s)} \tag{1.2}$$

Tại nhiệt độ phòng tiêu chuẩn 20°C, vận tốc âm thanh trong không khí là c ≈ 343,4 m/s (thường làm tròn thành 343 m/s).

Bảng 1.1 Vận tốc truyền âm trong các môi trường vật chất

Môi trường vật chất	Nhiệt độ	Vận tốc c (m/s)	Ghi chú
Không khí khô (chuẩn)	0°C	331,3	Điều kiện thời tiết lạnh, mùa đông vùng cao
Không khí phòng thu chuẩn	20°C	343,4	Giá trị thường dùng khi tính toán độ trễ
Không khí mùa hè	35°C	352,5	Ngoài trời nắng nóng; vận tốc tăng làm thay đổi độ trễ
Khí Heli (Helium)	20°C	1.007	Hiệu ứng 'giọng vịt Donald' — vận tốc lớn dịch chuyển tần số formant
Nước ngọt	25°C	1.497	Nhanh hơn không khí ~4,4 lần; thu tiếng thủy âm bằng Hydrophone
Nước biển (độ mặn 3,5%)	25°C	1.531	Độ mặn làm tăng tỷ trọng và vận tốc truyền âm
Gỗ thông, gỗ sồi	20°C	3.300 – 4.000	Vật liệu cộng hưởng nhạc cụ, sàn phòng, vách tiêu âm
Bê tông cốt thép	20°C	3.000 – 4.000	Lan truyền rung chấn cơ học qua kết cấu tường

Môi trường vật chất	Nhiệt độ	Vận tốc c (m/s)	Ghi chú
Thép kết cấu	20°C	5.120	Tiếng va đập kim loại truyền qua kết cấu thép rất nhanh
Chân không vũ trụ	–	0	Không tồn tại môi trường truyền sóng cơ học; hoàn toàn im lặng

Trở kháng âm học đặc trưng (Characteristic Acoustic Impedance, ký hiệu Z) là đại lượng vật lý quyết định mức độ sóng âm truyền qua hay bị phản xạ trở lại tại mặt phân cách giữa hai môi trường đàn hồi:

$$Z = \rho c \quad (\text{rayl} = \text{Pa·s/m}) \tag{1.3}$$

Trong đó: ρ là khối lượng riêng của môi trường (kg/m³) và c là vận tốc truyền âm trong môi trường đó. Không khí ở 20°C có Z ≈ 413 rayl (với ρ = 1,204 kg/m³ và c = 343 m/s), trong khi nước có Z ≈ 1,48 × 10⁶ rayl, chênh lệch gấp hơn 3.500 lần. Sự bất tương hợp trở kháng lớn này giải thích vì sao khi một người lặn xuống nước, họ hầu như không nghe thấy tiếng nói của người trên bờ: hơn 99,9% năng lượng sóng âm bị phản xạ ngược lại tại mặt phân cách không khí – nước. Nguyên lý vật lý này cũng lý giải vì sao chuỗi xương tai giữa của con người phải hoạt động như một hệ đòn bẩy cơ học để phối hợp trở kháng giữa không khí ống tai ngoài và dịch nội ốc tai trong, và vì sao một bức tường bê tông đặc dày cách âm tốt hơn nhiều so với một tấm mút xốp nhẹ.

VÍ DỤ 1.1 — TÍNH TOÁN VẬN TỐC ÂM VÀ ĐỘ TRỄ LOA NGOÀI TRỜI

Tình huống: Một buổi ca nhạc ngoài trời diễn ra lúc 14 giờ chiều, nhiệt độ môi trường đo được là 34°C. Người phụ trách âm thanh bố trí một cặp loa tăng cường (loa fill) cách sân khấu chính một khoảng cách d = 12 m và cần thiết lập độ trễ (delay) trên bàn mixer kỹ thuật số. Yêu cầu: tính vận tốc âm thanh thực tế và thời gian trễ cần cài đặt.

Hướng dẫn giải: Áp dụng công thức thực nghiệm nhiệt độ: c ≈ 331,3 + 0,606 × 34 = 331,3 + 20,60 = 351,9 m/s ≈ 352 m/s.

Thời gian truyền âm: t = d / c = 12 / 351,9 ≈ 0,0341 giây = 34,1 ms.

Quyết định kỹ thuật: Loa fill cần được cài đặt thời gian trễ 34 ms trên bộ xử lý tín hiệu (DSP) để sóng âm từ loa fill đến tai người nghe đồng pha hoặc chậm hơn một chút so với âm trực tiếp phát ra từ sân khấu. Nếu giữ nguyên giá trị mặc định ở 20°C (c = 343 m/s), thời gian trễ tính ra là 35,0 ms, sai lệch 0,9 ms; mức sai lệch này nhỏ ở cự ly 12 m nhưng tăng tỷ lệ theo khoảng cách khi loa đặt ở cự ly xa hơn.

1.1.4 Bước sóng trong không gian vật lý và ý nghĩa thiết kế loa siêu trầm

Bước sóng λ (wavelength) là quãng đường không gian mà sóng âm truyền đi được trong khoảng thời gian bằng một chu kỳ dao động toàn phần T. Mối liên hệ toán học giữa bước sóng, vận tốc âm và tần số được xác định như sau:

$$\lambda = \frac{c}{f} = cT \quad (\text{m}) \tag{1.4}$$

Bước sóng là cầu nối trực tiếp giữa các con số tần số trừu tượng trên màn hình máy tính với kích thước vật lý ngoài đời thực của căn phòng và thiết bị âm thanh:

Bảng 1.2 Bước sóng vật lý của âm thanh và Hệ quả kỹ thuật trong thiết kế âm học

Tần số f (Hz)	Bước sóng λ (ở 20°C, c = 343 m/s)	Hệ quả kỹ thuật
20 Hz (Siêu trầm)	17,15 m	Lớn hơn hầu hết phòng dựng thông thường; phòng trở nên 'trong suốt' với mode bậc 1
50 Hz (Sub-bass)	6,86 m	Tương đương chiều dài một căn phòng lớn; dễ tạo sóng đứng cộng hưởng mạnh
100 Hz (Bass)	3,43 m	Tương đương kích thước một phòng nhỏ; dễ gây ù rền
343 Hz (Mốc nhớ)	1,00 m	Mốc dễ nhớ: tần số 343 Hz tương ứng bước sóng 1 mét
1.000 Hz (1 kHz)	34,3 cm	Tương đương kích thước đầu người; mốc chuyển giao cơ chế định vị không gian ITD sang ILD
4.000 Hz (4 kHz)	8,58 cm	Tương đương đường kính họng kèn tweeter; âm thanh bắt đầu bức xạ có tính định hướng hẹp
10.000 Hz (10 kHz)	3,43 cm	Tương đương kích thước lưới chắn đầu micro (grille); dễ gây nhiễu xạ làm méo đáp tuyến
20.000 Hz (Treble)	1,71 cm	Tương đương độ dày tấm mút xốp mỏng; chỉ có thể tiêu âm hiệu quả dải tần số này

Nguyên lý cơ bản của âm học vật lý cho biết một vật cản hay một tấm vật liệu chỉ chắn hoặc hấp thụ hiệu quả sóng âm khi kích thước của nó lớn hơn ít nhất một nửa bước sóng (λ/2). Vì vậy, một tấm vách ngăn di động cao 1,2 m chỉ ngăn được tiếng ồn dải trung và cao, còn tiếng máy phát điện hoặc điều hòa ở 50–60 Hz (bước sóng 6–7 m) nhiễu xạ vòng qua vách mà không bị cản trở đáng kể. Tương tự, một tấm foam tiêu âm dán tường dày 5 cm chỉ bắt đầu hấp thụ từ khoảng 1,7 kHz trở lên; việc dán lớp mút mỏng lên tường phòng dựng để xử lý hiện tượng 'ùng ục' ở dải trầm là sai lầm điển hình do chưa nắm vững quan hệ giữa kích thước vật liệu và bước sóng.

Ý nghĩa đối với thiết kế loa Subwoofer: để bức xạ hiệu quả các tần số siêu trầm có bước sóng dài hàng chục mét, loa siêu trầm (Subwoofer) phải có màng loa đường kính lớn (15 đến 18 inch, tức 38–46 cm) cùng thùng loa thể tích lớn để tạo ra sự dịch chuyển thể tích không khí tương ứng. Ngược lại, loa tép (tweeter) bức xạ bước sóng chỉ vài milimét nên màng loa nhỏ, cỡ 1 inch, là đủ để bảo đảm góc phủ âm rộng.

1.1.5 Dạng sóng và mô hình lan truyền sóng cầu, sóng phẳng

Cách thức sóng âm lan tỏa trong không gian quyết định kỹ thuật đặt micro và việc tính toán mức suy hao âm lượng:

1. **Mô hình sóng cầu (Spherical wave, gần nguồn):** Khi kích thước vật lý của nguồn âm nhỏ hơn nhiều so với bước sóng bức xạ (như miệng người nói hay một nhạc cụ độc tấu), nguồn âm được mô hình hóa thành nguồn điểm (point source). Năng lượng âm bức xạ đồng đều ra mọi hướng trong không gian tự do, và mặt sóng tạo thành các mặt cầu đồng tâm lớn dần. Diện tích mặt cầu tăng theo bình phương bán kính ($S = 4\pi r^2$) nên năng lượng phân tán trên diện tích lớn hơn, và mức áp suất âm suy giảm 6 dB mỗi khi khoảng cách nhân đôi (định luật bình phương nghịch đảo).
2. **Mô hình sóng phẳng (Plane wave, xa nguồn):** Khi sóng cầu truyền đi một khoảng cách đủ xa so với nguồn phát, bán kính mặt cầu trở nên rất lớn và độ cong của mặt cầu giảm dần, do đó trên một diện tích khảo sát cục bộ như màng micro, mặt sóng có thể coi gần đúng là một mặt phẳng hình học. Trong trường sóng phẳng, các tia sóng song song với nhau và mức áp suất âm hầu như không suy giảm theo diện tích tiết diện ngang.

Ý nghĩa kỹ thuật: các microphone đo kiểm âm học chuyên dụng và micro thu âm đều được nhà sản xuất hiệu chuẩn đáp tuyến tần số trong trường sóng phẳng (Free-field hoặc Diffuse-field), trong khi hiệu ứng cận âm (Proximity Effect, tức sự tăng mạnh thành phần bass khi đặt micro sát miệng) chỉ xuất hiện trong trường sóng cầu ở cự ly rất gần nguồn.

1.2 Bốn đại lượng đặc trưng của sóng âm

Sóng sin (sine wave) là dạng dao động cơ học điều hòa thuần khiết và cơ bản nhất, chỉ chứa đúng một tần số duy nhất mà không có thành phần hài âm nào. Phương trình tổng quát của một sóng áp suất âm hình sin truyền trong không gian được biểu diễn như sau:

$$p(t) = P_{\max} \sin(2\pi ft + \varphi) \tag{1.5}$$

Trong đó: p(t) là áp suất âm tức thời tại thời điểm t (Pa); P_{max} là biên độ áp suất đỉnh (Pa); f là tần số dao động (Hz); $\omega = 2\pi f$ là tần số góc (rad/s); φ là góc pha ban đầu (radian). Theo định lý biến đổi Fourier (Fourier Transform), mọi âm thanh, dù phức tạp đến đâu, từ tiếng nổ, tiếng đàn violin cho đến lời nói thầm, đều có thể phân tích chính xác thành tổng của nhiều thành phần sóng sin đơn tần có biên độ, tần số và pha xác định. Vì vậy, kỹ thuật xử lý âm thanh xoay quanh bốn đại lượng cơ bản sau:

1.2.1 Biên độ, giá trị hiệu dụng RMS và nguy cơ méo cắt đỉnh

Biên độ đặc trưng cho độ lớn của dao động áp suất so với trạng thái áp suất khí quyển tĩnh. Về mặt sinh lý thính giác, biên độ càng lớn thì âm thanh nghe càng to và biên độ càng nhỏ thì âm thanh nghe càng bé. Trong kỹ thuật audio, quá trình chuyển đổi từ sóng âm sang tín hiệu điện tại đầu ra micro tạo nên điện áp tỷ lệ

thuận với áp suất âm. Để điều chỉnh chính xác mức tín hiệu trên bàn trộn (Mixer/DAW), người làm âm thanh cần phân biệt ba cách đo lường biên độ:

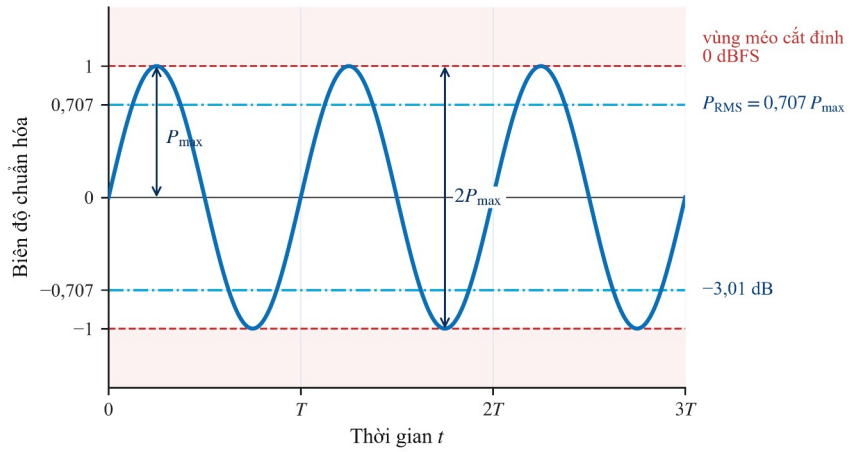
1. **Biên độ đỉnh (Peak Amplitude, A_{peak}):** là giá trị tức thời tuyệt đối lớn nhất của tín hiệu trong khoảng thời gian khảo sát: $A_{peak} = \max(|p(t)|)$. Đây là thông số giới hạn an toàn vật lý của hệ thống. Trong miền kỹ thuật số (Digital Audio), nếu biên độ đỉnh vượt quá ngưỡng toàn thang 0 dBFS, mạch chuyển đổi số (ADC/DAC) sẽ cắt phẳng đỉnh sóng sin và gây ra méo cắt đỉnh (Clipping). Méo cắt đỉnh tạo ra một chuỗi hài bậc lẻ rất chói tai, làm rách tiếng, vỡ tiếng và phá hỏng bản thu [18].
2. **Biên độ đỉnh – đỉnh (Peak-to-Peak Amplitude, $A_{p\text{-}p}$):** là khoảng cách biên độ toàn phần giữa đỉnh dương cực đại và đáy âm cực tiểu: $A_{p\text{-}p} = 2 \cdot A_{peak}$ (đối với sóng hình sin đối xứng). Đại lượng này thường được theo dõi trên máy hiện sóng (Oscilloscope) để kiểm tra độ dịch chuyển cơ học tối đa của màng loa công suất lớn và màng rung micro.
3. **Giá trị hiệu dụng (Root Mean Square, RMS):** là căn bậc hai của giá trị trung bình bình phương các giá trị tức thời trong một chu kỳ tuần hoàn T. Giá trị RMS phản ánh trực tiếp năng lượng và công suất hiệu dụng của sóng âm:

$$P_{\text{RMS}} = \sqrt{\frac{1}{T}\int_0^T p^2(t)\,dt} = \frac{P_{\max}}{\sqrt{2}} \approx 0{,}707\,P_{\max} \tag{1.6}$$

Ý nghĩa thực tiễn: màng nhĩ và não bộ con người có quán tính sinh lý nhất định nên không phản ứng tức thời với các đỉnh nhọn kéo dài mili-giây (Peak), mà cảm nhận độ to (loudness) dựa trên năng lượng trung bình tích lũy theo thời gian (RMS hoặc chuẩn đo LUFS). Do đó, một phát súng có đỉnh Peak rất cao nhưng thời lượng rất ngắn có thể nghe không to bằng tiếng động cơ xe tải có RMS cao và liên tục. Để định lượng tỷ lệ giữa đỉnh tức thời và năng lượng trung bình, kỹ thuật audio định nghĩa hệ số đỉnh (Crest Factor):

$$\text{CF} = 20\log_{10}\frac{P_{\max}}{P_{\text{RMS}}} \quad \text{(dB)} \tag{1.7}$$

Sóng sin thuần khiết có Crest Factor lý tưởng là $20\cdot\log_{10}(\sqrt{2}) \approx 3{,}01$ dB. Bản ghi âm dàn nhạc giao hưởng thu tự nhiên thường có Crest Factor rất lớn, từ 18 đến 22 dB, cho cảm giác âm thanh thoáng và giàu độ động. Ngược lại, các bản nhạc pop thương mại hoặc quảng cáo bị nén quá mức qua bộ Limiter trong cuộc chiến độ to (Loudness War) có thể chỉ còn Crest Factor 6–8 dB; âm thanh khi đó nghe rất to nhưng bị dẹt, thiếu độ động và gây mỏi thính giác nhanh cho người nghe.

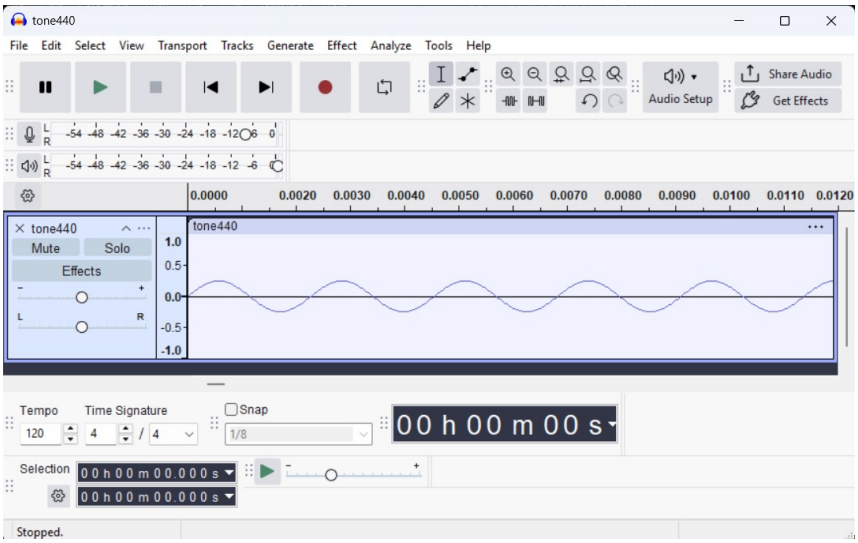


Hình 1.2 Biên độ đỉnh, đỉnh – đỉnh, giá trị hiệu dụng RMS và ngưỡng méo cắt đỉnh của hệ thống

1.2.2 Tần số, cao độ và thang đo logarit

Tần số f là số chu kỳ dao động toàn phần mà nguồn âm và phần tử môi trường thực hiện được trong 1 giây, đo bằng đơn vị Hertz (Hz), với chu kỳ dao động T = 1/f (giây). Tần số là đại lượng vật lý khách quan, còn cao độ (pitch) là cảm nhận chủ quan của tai người. Dao động càng nhanh thì tần số càng cao và âm thanh nghe càng thanh, cao (treble); ngược lại, dao động càng chậm thì âm phát ra càng trầm và nặng (bass).

Hình 1.3 là dạng sóng của một âm sin 440 Hz được phóng to trong Audacity cho tới khi thấy từng chu kỳ: trong 12 ms có hơn năm chu kỳ, khớp với chu kỳ T = 1/440 s ≈ 2,27 ms.



Hình 1.3 Dạng sóng âm sin 440 Hz phóng to đến mức thấy từng chu kỳ, kèm thước thời gian (Audacity)

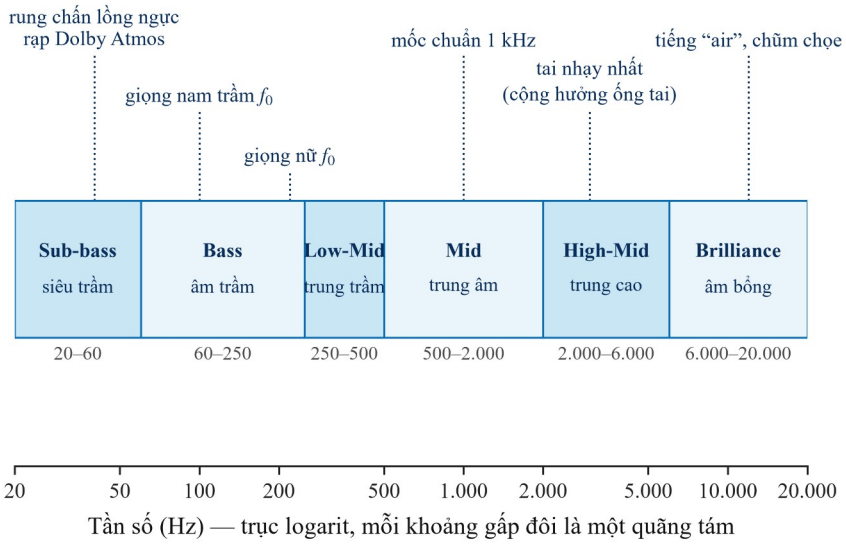
Nguồn: ảnh chụp màn hình do người soạn thực hiện (Audacity 3.7.7), 05/11/2026.

Dải thính giác bình thường của con người cảm nhận được tần số từ 16 Hz đến 20.000 Hz (20 kHz). Các sóng cơ học dưới 16 Hz gọi là hạ âm (infrasound), còn trên 20.000 Hz gọi là siêu âm (ultrasound). Cảm nhận về cao độ của tai người tuân theo thang logarit chứ không theo thang tuyến tính: mỗi khi tần số tăng gấp đôi, cao độ cảm nhận tăng đúng một quãng tám (Octave). Do đó, dải tần nghe được

của con người trải dài đúng 10 quãng tám (20 Hz → 40 Hz → 80 Hz → 160 Hz → 320 Hz → 640 Hz → 1,28 kHz → 2,56 kHz → 5,12 kHz → 10,24 kHz → 20,48 kHz). Decade là khoảng tần số tăng gấp 10 lần (ví dụ từ 100 Hz lên 1.000 Hz) và thường được dùng làm trục hoành của đồ thị máy phân tích phổ (RTA). Tần số nốt La giữa chuẩn quốc tế để so dây nhạc cụ và căn chuẩn thiết bị phòng thu là A4 = 440 Hz (một số dàn nhạc giao hưởng châu Âu lấy chuẩn 442 Hz).

Bảng 1.3 Sáu phân vùng tần số thường dùng trong xử lý âm thanh

Phân vùng tần số	Dải tần số vật lý	Đặc tính cảm nhận và ứng dụng
1. Sub-bass (Siêu trầm)	20 Hz – 60 Hz	Cảm nhận chủ yếu qua rung chấn cơ học ở lồng ngực người nghe (loa Subwoofer); tiếng nổ, sấm sét. Cần cắt bỏ bằng bộ lọc High-Pass (80 Hz) trên track thoại để tránh ùng.
2. Bass (Âm trầm)	60 Hz – 250 Hz	Nền tảng ấm áp và uy lực của giọng nam, tiếng đàn cello, kick drum, động cơ xe hơi. Là vùng tần số dễ bị tích tụ năng lượng gây đục ngầu (muddy) bản hòa âm nếu nhiều nhạc cụ cùng chiếm chỗ.
3. Low-Mid (Trung trầm)	250 Hz – 500 Hz	Quyết định độ dày và thân hình ('body') của lời thoại và nhạc cụ. Nếu dư thừa dải này, âm thanh sẽ có cảm giác 'nghẹt mũi', 'hộp các-tông' như đang nói trong chum vại.
4. Mid (Trung âm)	500 Hz – 2.000 Hz	Vùng chứa phần lớn năng lượng cơ bản của tiếng nói con người và tiếng động sinh hoạt đời thường; quyết định trực tiếp tới việc nghe hiểu nội dung lời nói.
5. High-Mid (Trung cao)	2.000 Hz – 6.000 Hz	Vùng thính giác con người nhạy cảm nhất (đặc biệt là 3–4 kHz do cộng hưởng ống tai). Quyết định độ rõ nét câu từ (Speech Intelligibility) và tiếng kim loại đanh thép. Tăng quá tay gây chói gắt và mệt tai rất nhanh.
6. Brilliance / Treble	6.000 Hz – 20.000 Hz	Độ trong trẻo, không gian khoáng đạt ('air'), tiếng thì thào, tiếng xào xạc của lá cây, tiếng cymbal. Là dải tần bị suy thoái sớm nhất theo tuổi tác của con người.



Hình 1.4 Sáu phân vùng tần số của dải nghe được và các mốc thường dùng trong xử lý âm thanh

Bài thực hành 1.1 — Tạo và xuất tệp sóng sin chuẩn bằng Python

Đoạn mã nguồn dưới đây minh họa các bước lập trình cơ bản để tạo một mảng số thực biểu diễn sóng sin thuần khiết và lưu dưới dạng tệp âm thanh chuẩn WAV 16-bit PCM:

Liệt kê 1.1 src/hello.py — Tạo và xuất sóng sin đơn âm ra file WAV

```
# src/hello.py - Chương trình tổng hợp sóng sin cơ bản
import numpy as np
import soundfile as sf

# 1. Thiết lập các tham số vật lý của tín hiệu
SAMPLE_RATE = 44100   # Tần số lấy mẫu tiêu chuẩn CD (Hz)
DURATION = 2.0        # Độ dài thời gian của tín hiệu (giây)
FREQUENCY = 440.0     # Tần số dao động sóng sin nốt A4 (Hz)
AMPLITUDE = 0.5       # Biên độ đỉnh (tương đương -6 dBFS)

# 2. Tạo trục thời gian rời rạc t = [0, 1/fs, 2/fs, ..., (N-1)/fs]
total_samples = int(SAMPLE_RATE * DURATION)
t = np.linspace(0, DURATION, total_samples, endpoint=False)

# 3. Tính toán công thức toán học sóng sin trên từng mẫu: y = A *
sin(2*pi*f*t)
audio_data = AMPLITUDE * np.sin(2.0 * np.pi * FREQUENCY * t)

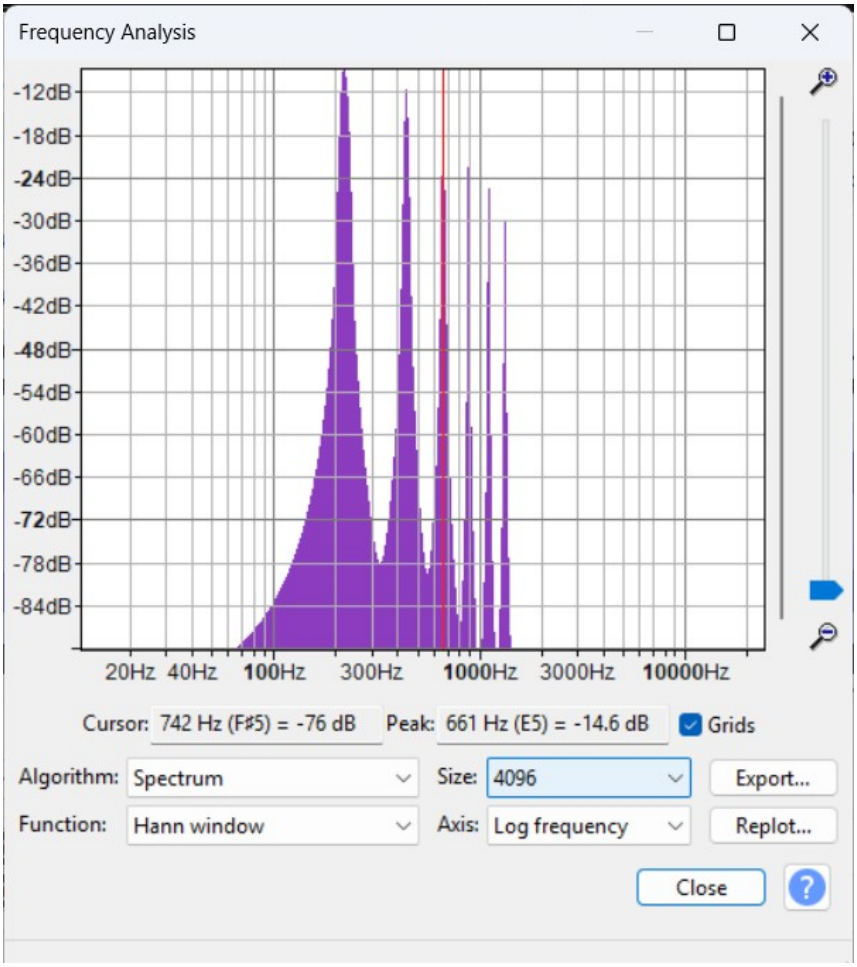
# 4. Xuất dữ liệu mảng numpy ra file âm thanh WAV chuẩn PCM 16-bit
output_filename = "sine_440hz.wav"
sf.write(output_filename, audio_data, SAMPLE_RATE, subtype='PCM_16')
print(f"[THÀNH CÔNG] Đã tạo file: {output_filename}")
print(f"Tổng số mẫu dữ liệu: {len(audio_data)} samples | Dải giá trị:
[{audio_data.min():.2f}, {audio_data.max():.2f}]")
```

1.2.3 Âm sắc, chuỗi hài Fourier và mô hình đường bao ADSR

Khi hai người cùng đọc một câu ở cùng cao độ nốt La 440 Hz và cùng âm lượng, người nghe vẫn phân biệt được chất giọng ấm của người này với giọng đanh của người kia; tương tự, cây đàn violin và cây kèn trumpet cùng chơi nốt A4 = 440

Hz vẫn mang màu sắc âm thanh khác biệt rõ rệt. Sự khác biệt này do *âm sắc* (Timbre) quyết định, đại lượng được ví như 'dấu vân tay âm thanh' của từng nguồn.

Hình 1.5 là phổ của một âm phức hợp gồm sáu hài của cơ bản 220 Hz, đo bằng chức năng Plot Spectrum của Audacity với cửa sổ Hann và cỡ FFT 4096; các đỉnh nằm đúng ở 220, 440, 660, 880 Hz… với biên độ giảm dần.



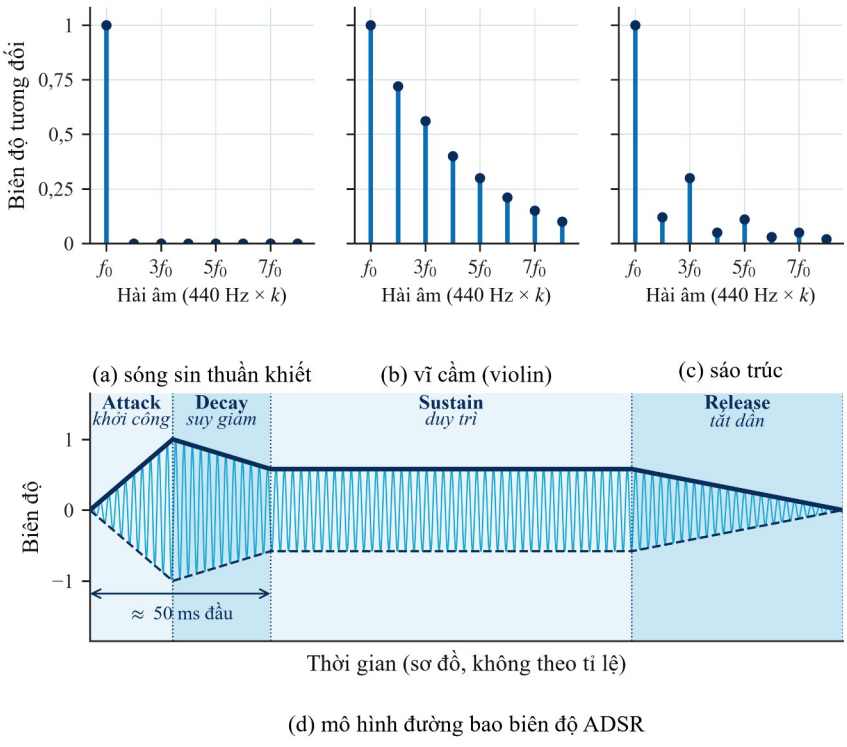
Hình 1.5 Phổ vạch hài của một âm phức hợp (cơ bản 220 Hz) trong hộp thoại Frequency Analysis của Audacity

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Audacity 3.7.7), 05/11/2026.

Âm sắc được cấu thành từ hai yếu tố cơ bản là phổ hài tần số (Harmonic Spectrum) và đường bao biên độ theo thời gian (ADSR Envelope) [15]:

1. **Phổ hài Fourier (Harmonics):** Khi một nguồn âm tự nhiên dao động, ngoài tần số cơ bản f_0 (fundamental frequency) quyết định cao độ, nguồn âm luôn phát sinh kèm theo một chuỗi tần số cộng hưởng gọi là hài âm (harmonics) có tần số là bội số nguyên của f_0 ($2f_0$, $3f_0$, $4f_0$, $5f_0$...). Tỷ lệ biên độ tương đối giữa các hài âm này tạo nên phổ hài đặc trưng:
2. **Hài bậc chẵn (Even Harmonics):** Các bội số $2f_0$, $4f_0$, $6f_0$... tạo cảm giác âm thanh ấm, tròn và dày. Đây là lý do chất âm của tiền khuếch đại bóng đèn điện tử (tube preamp) hoặc máy ghi băng từ (tape saturation) thường được dùng để làm ấm, làm mềm giọng nói.
3. **Hài bậc lẻ (Odd Harmonics):** Các bội số $3f_0$, $5f_0$, $7f_0$... mang tính chất sắc, gắt và kim loại, dễ gây cảm giác căng thẳng cho người nghe.

4. **Mô hình đường bao biên độ 4 giai đoạn ADSR:** Bên cạnh phổ tần số tĩnh, sự biến thiên biên độ theo thời gian đóng vai trò quyết định trong việc nhận diện nguồn âm và được mô hình hóa qua 4 giai đoạn ADSR kinh điển:
5. **Attack (Khởi phát):** Thời gian từ khi âm thanh bắt đầu phát sinh đến khi đạt biên độ đỉnh cực đại (từ vài mili-giây đến vài trăm ms). 50 mili-giây đầu tiên của giai đoạn Attack chứa hơn 80% thông tin nhận dạng đặc tính vật lý của nguồn âm. Nếu cắt bỏ 50 ms đầu tiên của nốt đàn piano hoặc tiếng va đập đồ vật, người nghe sẽ khó nhận ra đó là âm thanh gì. Do đó, việc đặt thời gian Attack của bộ nén (compressor) quá nhanh sẽ làm dẹt tiếng nổ và tiếng trống, đồng thời làm mất độ nảy (transient) của âm thanh.
6. **Decay (Suy giảm ban đầu):** Thời gian biên độ giảm từ mức đỉnh Attack xuống mức duy trì ổn định.
7. **Sustain (Duy trì):** Mức biên độ duy trì ổn định trong suốt khoảng thời gian nguồn kích thích tiếp tục tác động (khi người nói ngân dài giọng, luồng gió thổi đều hoặc kéo vĩ đàn violin).
8. **Release (Tắt dần):** Thời gian cần thiết để năng lượng âm thanh tắt hẳn về trạng thái im lặng sau khi nguồn kích thích cơ học ngừng hoàn toàn. Giai đoạn Release ảnh hưởng đến độ tự nhiên và cảm nhận về không gian của âm thanh.



Hình 1.6 Âm sắc: phổ hài Fourier của nốt La 440 Hz ở ba nguồn âm khác nhau và mô hình đường bao biên độ ADSR

Bài thực hành 1.2 — Tổng hợp âm phức hợp đa hài và chuẩn hóa biên độ

Để tạo ra âm thanh giàu âm sắc như một nhạc cụ thực thụ, người đọc áp dụng nguyên lý tổng hợp cộng (Additive Synthesis) bằng cách cộng dồn các sóng sin họa âm bậc cao. Đoạn mã dưới đây minh họa kỹ thuật tổng hợp nốt A4 gồm 4

họa âm, kèm thuật toán chuẩn hóa đỉnh (Peak Normalization) nhằm triệt tiêu nguy cơ méo cắt đỉnh (clipping):

Liệt kê 1.2 src/rich_sound.py — Tổng hợp đa hài Fourier và thuật toán chuẩn hóa biên độ

```
# src/rich_sound.py - Tổng hợp chuỗi hài Fourier và Chuẩn hóa chống méo
Clipping
import numpy as np
import soundfile as sf

SAMPLE_RATE = 44100
DURATION = 2.5
F0 = 440.0  # Tần số cơ bản nốt A4

t = np.linspace(0, DURATION, int(SAMPLE_RATE * DURATION), endpoint=False)

# Cấu hình chuỗi hài: (Hệ số tần số k, Biên độ tương đối)
# Âm thanh sáng và ấm giống tiếng kèn đồng hoặc dây đàn
harmonics = [
    (1, 1.00),   # Họa âm cơ bản f0 = 440 Hz (biên độ mạnh nhất)
    (2, 0.50),   # Họa âm bậc hai 2f0 = 880 Hz (quãng 8)
    (3, 0.35),   # Họa âm bậc ba 3f0 = 1320 Hz (quãng 5 của quãng 8)
    (4, 0.20),   # Họa âm bậc bốn 4f0 = 1760 Hz (quãng 8 thứ hai)
]

# Khởi tạo mảng tín hiệu rỗng chứa toàn số 0
rich_sound = np.zeros_like(t)

# Thực hiện phép cộng dồn chuỗi Fourier
for k, amp in harmonics:
    harmonic_freq = F0 * k
    rich_sound += amp * np.sin(2.0 * np.pi * harmonic_freq * t)

# KIỂM TRA BIÊN ĐỘ ĐỈNH VÀ CHUẨN HÓA (PEAK NORMALIZATION)
peak_val = np.max(np.abs(rich_sound))
print(f"Biên độ đỉnh thô trước khi chuẩn hóa: {peak_val:.3f}")

# Ngưỡng biên độ mong muốn: -1 dBFS tương đương 10^(-1/20) ≈ 0.891
TARGET_PEAK = 0.891

if peak_val > TARGET_PEAK:
    print("Cảnh báo: tín hiệu vượt ngưỡng! Tiến hành chuẩn hóa để chống
Clipping...")
    rich_sound = (rich_sound / peak_val) * TARGET_PEAK

# Xuất ra file âm thanh chất lượng cao
output_file = "rich_sound_440hz.wav"
sf.write(output_file, rich_sound, SAMPLE_RATE, subtype='PCM_16')
print(f"[THÀNH CÔNG] Đã lưu tín hiệu đa hài vào: {output_file}")
print(f"Biên độ đỉnh sau chuẩn hóa: {np.max(np.abs(rich_sound)):.3f} (An toàn
dưới 0 dBFS)")
```

1.2.4 Góc pha, hiện tượng triệt tiêu và hiệu ứng lọc lược

Góc pha φ đặc trưng cho trạng thái và vị trí tức thời của sóng âm trong chu kỳ dao động tại một thời điểm xác định, đo bằng độ (0° đến 360°) hoặc radian. Sự

tương tác pha giữa hai sóng âm cùng tần số là vấn đề then chốt khi thu âm bằng nhiều micro [16], [14]:

1. **Hiện tượng cùng pha (In-Phase, Δφ = 0°):** Hai sóng âm có đỉnh và đáy hoàn toàn trùng khớp nhau về thời gian, nên khi cộng lại, biên độ sóng tổng hợp tăng gấp đôi ($A_{tổng} = 2 \cdot A$), tương ứng mức áp suất âm thanh tăng thêm đúng +6 dB.
2. **Hiện tượng ngược pha (Out-of-Phase, Δφ = 180° hay π rad):** Khi một sóng đạt đỉnh dương cực đại thì sóng kia đạt đỉnh âm đối xứng, nên năng lượng dao động áp suất của hai sóng triệt tiêu hoàn toàn lẫn nhau và biên độ tổng hợp bằng không (p(t) = 0 Pa, tương ứng -∞ dB).

Cần phân biệt hai khái niệm thường bị ngộ nhận là đảo cực (Polarity Inversion) và lệch pha (Phase Shift). Đảo cực là phép nhân tín hiệu với -1 trên toàn bộ các tần số (tác dụng của nút 'Ø' trên bàn mixer), trong khi lệch pha là kết quả của một độ trễ thời gian vật lý Δt, với độ lệch pha $\Delta\varphi = 2\pi \cdot f \cdot \Delta t$ phụ thuộc trực tiếp vào tần số f. Do đó, nút đảo cực Ø không thể khắc phục lỗi lệch pha do khoảng cách gây ra; chỉ có công cụ căn chỉnh thời gian (Time Alignment / Delay) mới xử lý được triệt để lỗi này.

Hiện tượng lọc lược (Comb Filtering) xảy ra khi hai bản sao của cùng một tín hiệu âm thanh được trộn lại với một độ trễ thời gian nhỏ Δt (thường do âm phản xạ từ mặt bàn kính, tường cứng, hoặc do thu một nguồn âm bằng hai micro ở hai khoảng cách khác nhau); kết quả là đáp tuyến tần số có vô số đỉnh cộng hưởng (+6 dB) xen kẽ các hõm triệt tiêu hoàn toàn (-∞ dB) đều đặn như răng lược. Các tần số bị triệt tiêu hoàn toàn (nulls) được tính chính xác bằng công thức:

$$f_{\text{null}} = \frac{2n + 1}{2\,\Delta t}, \qquad n = 0, 1, 2, 3, \ \ldots \tag{1.8}$$

Ví dụ: với độ trễ chỉ 1 ms (0,001 s), các hõm triệt tiêu rơi vào các tần số 500 Hz, 1.500 Hz, 2.500 Hz, 3.500 Hz... Hậu quả là âm thanh bị lọc lược trở nên mỏng và rỗng, nghe như phát ra từ trong một đường ống kim loại, làm suy giảm chất lượng giọng nói. Để hạn chế hiện tượng này khi thu âm bằng nhiều micro, người đọc cần nắm vững Quy tắc 3:1:

Quy tắc khoảng cách 3:1 khi bố trí nhiều micro

Nguyên lý định lượng: khi bố trí hai microphone độc lập để thu âm hai người đang nói chuyện trong cùng một không gian, khoảng cách giữa hai micro (dmic-mic) tối thiểu phải gấp 3 lần khoảng cách từ mỗi người nói đến micro của chính họ ($d_{nguồn\text{-}mic}$): $d_{mic\text{-}mic} \geq 3 \cdot d_{nguồn\text{-}mic}$.

Cơ sở vật lý: theo định luật bình phương nghịch đảo, khi khoảng cách tăng gấp 3 lần, tín hiệu âm thanh của người nói A lọt vào micro của người nói B bị suy giảm tự nhiên từ 9,5 dB đến 12 dB so với tín hiệu chính. Mức chênh lệch này đủ lớn để hiện tượng triệt tiêu pha và lọc lược khi trộn hai kênh trở nên không đáng kể (ảnh hưởng dưới 0,5 dB).

Giải pháp thu âm bàn talkshow / họp báo: khi người nói ngồi quanh một mặt bàn phẳng cứng, sóng trực tiếp từ miệng và sóng phản xạ từ mặt bàn vào micro tạo độ trễ vài trăm micro-giây, gây lọc lược nặng. Một giải pháp là sử dụng micro áp suất bề mặt (Boundary Microphone / PZM) gắn sát mặt bàn: khi đó âm trực tiếp và âm phản xạ chạm màng thu cùng lúc ($\Delta t \approx 0$) nên hiện tượng lọc lược bị triệt tiêu hoàn toàn.

THỰC HÀNH HIỆN TRƯỜNG 1.1 — QUAN SÁT PHASE CANCELLATION VÀ KIỂM TRA ĐỘ TƯƠNG THÍCH MONO

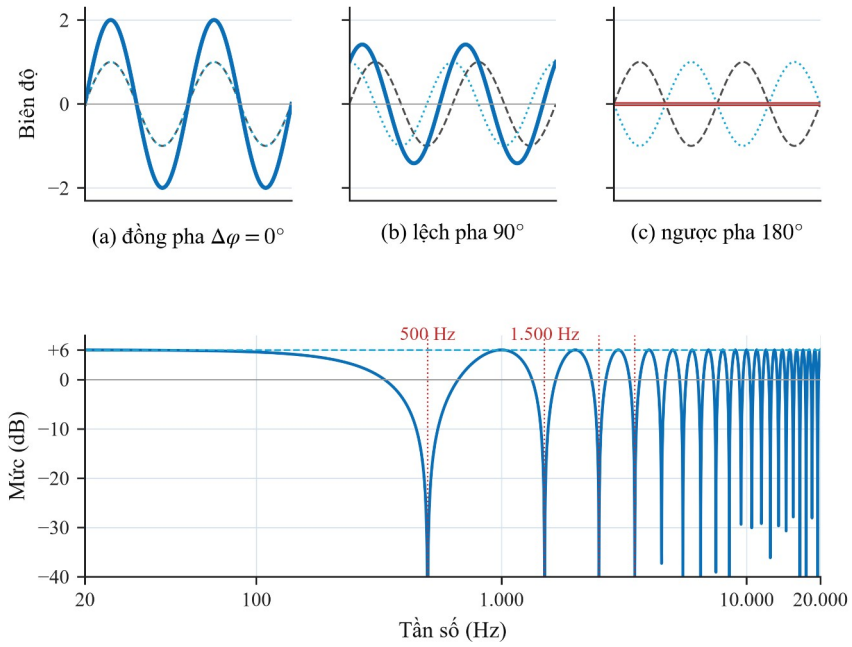
Quy trình thực hiện: trong phần mềm DAW (DaVinci Resolve Fairlight, Pro Tools hoặc Reaper), mở một dự án thu thanh trống, rồi cắm hai micro thu một nhạc cụ gõ (hoặc tiếng vỗ tay), một micro đặt gần (0,3 m) và một micro đặt xa (1,0 m). Ghi âm đồng thời và phát lại ở chế độ Stereo.

Thao tác kỹ thuật: 1. Chuyển nút điều khiển pan của cả hai track về chính giữa (trộn thành Mono) và lắng nghe âm sắc: tiếng trầm bị hụt hẳn và tiếng trung nghe rỗng do lọc lược.

2. Nhấn nút đảo cực pha (Ø / Phase Invert) trên một trong hai kênh: âm sắc thay đổi rõ rệt ngay lập tức.

3. Sử dụng công cụ phóng to mẫu sóng (Sample Nudge) để dịch chuyển track micro xa về phía trước sao cho trùng khớp đỉnh sóng đầu tiên của micro gần: âm thanh trở nên dày và đầy đặn trở lại.

Bài học nghề nghiệp: một bản mix nghe hay trên hệ thống loa kiểm âm Stereo của phòng dựng nhưng bị mất tiếng, mất lời khi phát trên điện thoại di động hoặc loa tivi thông thường là dấu hiệu điển hình của lỗi lệch pha. Vì vậy, kiểm tra độ tương thích Mono (Mono Compatibility) là bước cần thực hiện trước khi xuất tệp phát hành.



Hình 1.7 Tương tác pha giữa hai sóng âm và đáp tuyến lọc lược (Comb Filtering) khi trộn tín hiệu với bản sao trễ 1 ms

Bài thực hành 1.3 — Quan sát các đại lượng sóng và phổ trong Audacity

Mục tiêu bài thực hành: người đọc làm quen trực quan với việc quan sát sóng âm trên miền thời gian (Waveform) và miền tần số (Spectrogram), đồng thời kiểm chứng hiện tượng triệt tiêu pha bằng phần mềm nguồn mở Audacity.

Bước 1 — Sinh sóng sin chuẩn: Khởi động phần mềm Audacity, truy cập menu Generate → Tone... Chọn dạng sóng Sin (Sine), thiết lập tần số f = 440 Hz, biên độ Amplitude = 0.8, độ dài Duration = 3.0 giây, rồi nhấn Generate để tạo track âm thanh.

Bước 2 — Đo lường chu kỳ sóng: Sử dụng công cụ Zoom in (phím tắt Ctrl + 1) để phóng to dạng sóng, đếm chu kỳ dao động và xác minh bằng thước đo thời gian: chu kỳ của nốt A4 là T = 1 / 440 ≈ 0,00227 giây = 2,27 mili-giây.

Bước 3 — Phân tích phổ tần số: Nhấp chuột vào mũi tên bên cạnh tên track âm thanh ở bảng điều khiển bên trái và chọn Spectrogram. Người đọc quan sát thấy một đường kẻ ngang mảnh duy nhất và sáng rõ tại mốc tần số 440 Hz, minh chứng cho việc sóng sin chỉ chứa một thành phần tần số cơ bản, không có họa âm.

Bước 4 — Thử nghiệm triệt tiêu pha: Truy cập Tracks → Add New → Mono Track, rồi tạo một sóng sin thứ hai cũng ở tần số 440 Hz với cùng biên độ 0.8. Chọn toàn bộ track thứ hai, truy cập menu Effect → Invert (Đảo pha 180°), sau đó nhấn phím Space để phát đồng thời cả hai track: loa hoàn toàn im lặng, vì hai sóng âm cùng tần số và ngược pha đã triệt tiêu lẫn nhau.

Bước 5 — So sánh với sóng vuông: Tạo một tone sóng vuông 440 Hz và quan sát Spectrogram: xuất hiện thêm hàng loạt vạch sáng ở các tần số lẻ: 1.320 Hz (3f₀), 2.200 Hz (5f₀), 3.080 Hz (7f₀)... Đây là chuỗi hài Fourier, minh chứng sóng vuông được cấu thành từ tổng các sóng sin họa âm bậc lẻ.

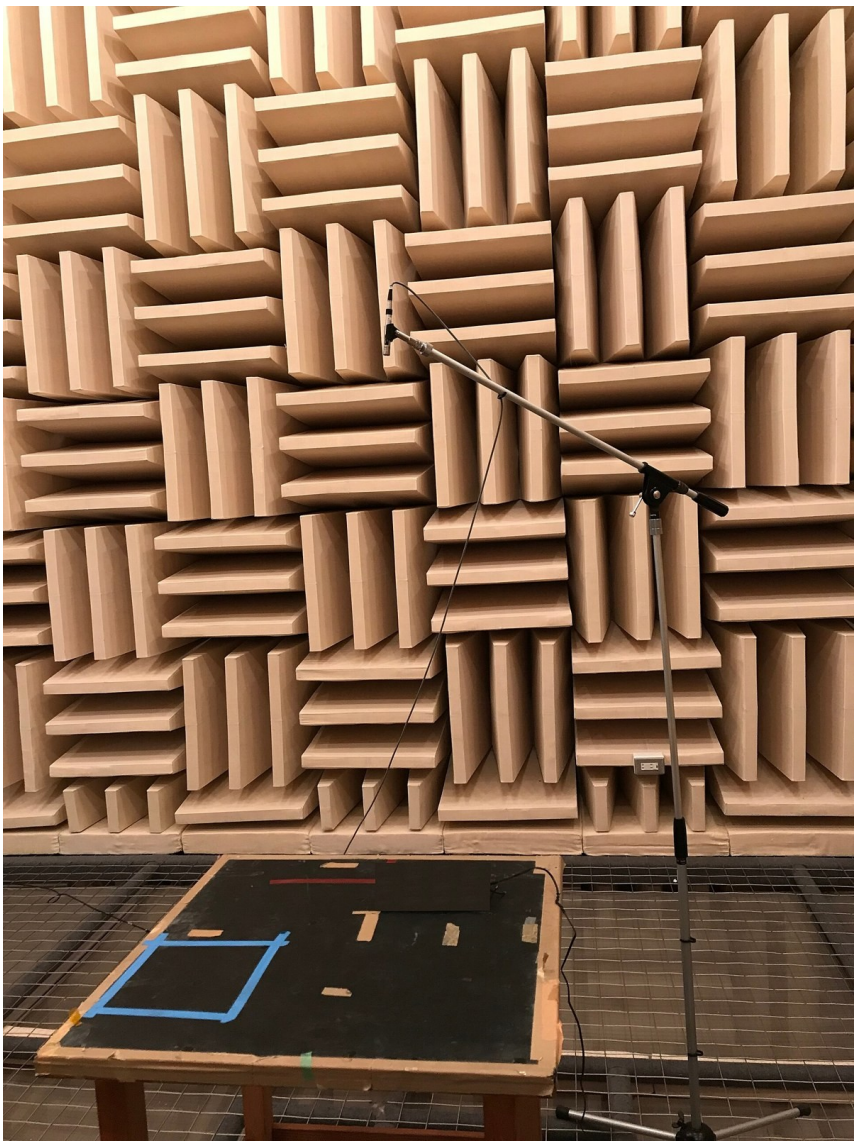
1.3 Âm học phòng

Khi một nguồn âm phát ra trong không gian kín, sóng âm va đập liên tục vào các bề mặt bao quanh (tường, trần, sàn, đồ đạc, cửa kính). Sự tương tác năng lượng phức tạp này tạo nên trường âm của căn phòng, ảnh hưởng trực tiếp đến chất lượng bản thu và độ chính xác của môi trường nghe kiểm tra.

1.3.1 Cơ chế phản xạ, dội âm và hệ số hấp thụ âm của vật liệu

Phản xạ âm học xảy ra khi sóng âm gặp bề mặt vật thể có trở kháng âm học Z lớn hơn nhiều so với không khí ($Z_{không_khí} \approx 413$ rayl). Trên mặt phẳng cứng nhẵn, sóng âm tuân theo định luật phản xạ hình học: góc phản xạ đúng bằng góc tới. Bề mặt lồi làm tán xạ năng lượng âm thanh ra nhiều hướng; ngược lại, bề mặt lõm (như trần vòm cong, góc tường thụt) làm hội tụ sóng âm vào một tiêu điểm cục bộ, gây méo trường âm nghiêm trọng trong phòng thu [8].

Hình 1.8 cho thấy một phòng đo vô hưởng, nơi các bề mặt được phủ khối nêm hấp thụ âm nhằm đưa hệ số hấp thụ về xấp xỉ giá trị lý tưởng và hầu như loại bỏ sóng phản xạ.



Hình 1.8 Phòng đo vô hưởng với các khối nêm hấp thụ âm phủ tường và micro đo trên giá đỡ

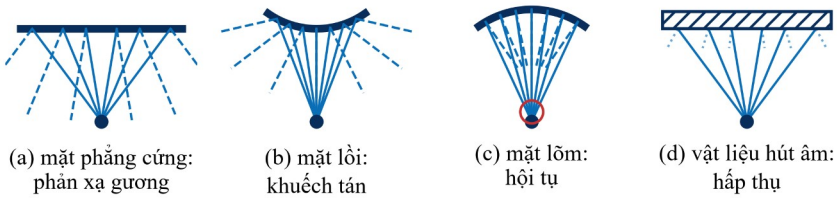
Nguồn: Togabi, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).

Khả năng tiêu hao năng lượng âm thanh của một vật liệu được định lượng thông qua hệ số hấp thụ âm thanh α (Absorption Coefficient), nhận giá trị không thứ nguyên từ 0 (phản xạ hoàn toàn 100% năng lượng) đến 1 (hấp thụ trọn vẹn 100% năng lượng). Hệ số α thay đổi mạnh theo tần số sóng âm:

Bảng 1.4 Hệ số hấp thụ âm thanh α của một số vật liệu thường gặp

Vật liệu bề mặt	α ở 125 Hz (Trầm)	α ở 500 Hz (Trung)	α ở 2.000 Hz (Cao)	Ý nghĩa ứng dụng và Lưu ý thực tiễn
Bê tông trần, gạch men lát sàn	0,01	0,02	0,02	Phản xạ 98% năng lượng; tạo dội âm hỗn loạn, tối kỵ trong phòng thu
Kính cửa sổ (dày 4 mm)	0,35	0,08	0,04	Hấp thụ nhẹ dải trầm do dao động uốn tấm kính; phản xạ toàn bộ dải cao

Vật liệu bề mặt	α ở 125 Hz (Trầm)	α ở 500 Hz (Trung)	α ở 2.000 Hz (Cao)	Ý nghĩa ứng dụng và Lưu ý thực tiễn
Thảm trải sàn nỉ dày	0,05	0,25	0,60	Chỉ hút âm tốt dải trung –cao; hoàn toàn không có tác dụng với tiếng trầm
Rèm nhung dày, xếp nếp 100%	0,15	0,55	0,70	Giải pháp xử lý dội âm cơ động, dễ lắp đặt và tháo dỡ
Tấm bông khoáng Rockwool 50 mm (60 kg/m³)	0,20	0,75	0,95	Tiêu âm phổ thông cho vách phòng thu thoại; hiệu quả từ 500 Hz trở lên
Bông khoáng 100 mm + khoang đệm khí 100 mm	0,55	0,95	0,98	Cấu hình tiêu âm tiêu chuẩn mở rộng hiệu quả xuống dải trầm sâu
Bẫy âm trầm cộng hưởng góc (Bass Trap)	0,80 – 0,90	0,30	0,15	Thiết bị bắt buộc tại các góc phòng để triệt tiêu sóng đứng dưới 150 Hz
Người ngồi trên ghế đệm nỉ	0,40	0,80	0,90	Phòng khi đông người có thời gian âm vang ngắn hơn khi trống



Hình 1.9 Bốn kiểu tương tác của sóng âm với bề mặt: phản xạ gương, khuếch tán, hội tụ và hấp thụ

1.3.2 Phân định sóng trực tiếp, tiếng vang và âm vang

Khi một nguồn âm phát ra trong không gian kín, tín hiệu thu nhận được tại màng micro hoặc tai người nghe bao gồm ba thành phần kế tiếp nhau theo trục thời gian:

1. **Sóng âm trực tiếp (Direct Sound):** là sóng âm lan truyền theo đường thẳng ngắn nhất từ nguồn âm tới màng micro mà không va chạm vào bất kỳ vật cản nào; sóng này bảo toàn nguyên vẹn âm sắc, mang năng lượng mạnh nhất và cung cấp thông tin định vị không gian chính xác.
2. **Tiếng vang (Echo):** xuất hiện khi sóng âm phản xạ từ một bề mặt cứng ở cự ly xa và dội về tai người nghe với độ trễ thời gian $\Delta t \geq 0{,}1$ giây (100 ms) so với sóng trực tiếp. Do quán tính lưu giữ âm thanh của não bộ chỉ kéo dài khoảng 50–100 ms, khi độ trễ vượt quá 100 ms, thính giác con người phân tách thành hai sự kiện âm thanh rời rạc. Về mặt vật lý, để xảy ra hiện tượng tiếng vang trong phòng kín ở 20°C (c = 343 m/s), quãng đường đi và về của

sóng phản xạ phải tối thiểu là: $s = c \cdot \Delta t = 343 \times 0{,}1 = 34{,}3$ m (tương ứng bức tường phản xạ phải cách nguồn âm ít nhất 17,2 mét).

3. **Âm vang (Reverberation):** xuất hiện khi các bề mặt phản xạ ở cự ly gần (dưới 17 mét), khiến hàng nghìn đợt sóng dội liên tiếp từ trần, sàn, tường trở về tai trong khoảng thời gian rất ngắn (độ trễ dưới 50 ms). Não bộ không thể phân tách các đợt sóng này thành các âm rời rạc mà hợp nhất chúng thành một 'đuôi âm vang' mịn, suy giảm dần sau khi nguồn âm ngừng phát. Âm vang tạo cảm giác về quy mô không gian (hang động sâu, nhà thờ lớn, hay phòng ngủ nhỏ); tuy nhiên, nếu âm vang của phòng quá lớn, đuôi âm vang của từ đi trước sẽ chèn lên từ đi sau và làm giảm độ rõ của lời nói (Speech Intelligibility).

1.3.3 Sóng đứng và vấn đề dội âm dải trầm tại các góc phòng

Sóng đứng (standing wave) là hiện tượng vật lý khó kiểm soát nhất trong âm học phòng thu và phòng dựng phim. Khi một sóng âm lan truyền tới bức tường cứng đối diện, nó bị phản xạ ngược trở lại; sóng tới và sóng phản xạ có cùng tần số, cùng bước sóng nên giao thoa liên tục với nhau, tạo nên một trường sóng tĩnh trong không gian [10]:

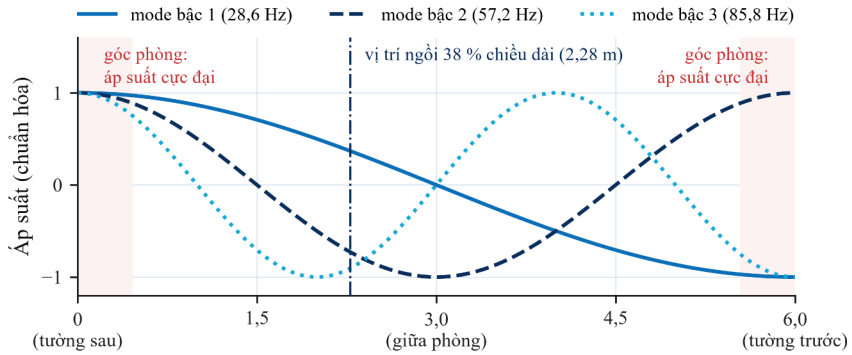
1. **Bụng sóng (Antinode — V):** là những tọa độ không gian nơi sóng tới và sóng phản xạ cùng pha, năng lượng dao động áp suất cộng hưởng lên cực đại. Tại các bề mặt tường cứng và đặc biệt là tại các góc phòng (nơi gặp nhau của 2 hoặc 3 mặt phẳng cứng), áp suất âm luôn đạt giá trị cực đại của bụng sóng.
2. **Nút sóng (Node — N):** là những tọa độ không gian nơi sóng tới và sóng phản xạ ngược pha 180°, năng lượng dao động áp suất bị triệt tiêu hoàn toàn về không. Khoảng cách giữa hai bụng sóng liên tiếp hoặc hai nút sóng liên tiếp đúng bằng nửa bước sóng ($\lambda/2$).

Tần số cộng hưởng sóng đứng của một căn phòng hình hộp chữ nhật có kích thước chiều dài (L), chiều rộng (W) và chiều cao (H) được xác định chính xác bằng công thức Rayleigh kinh điển:

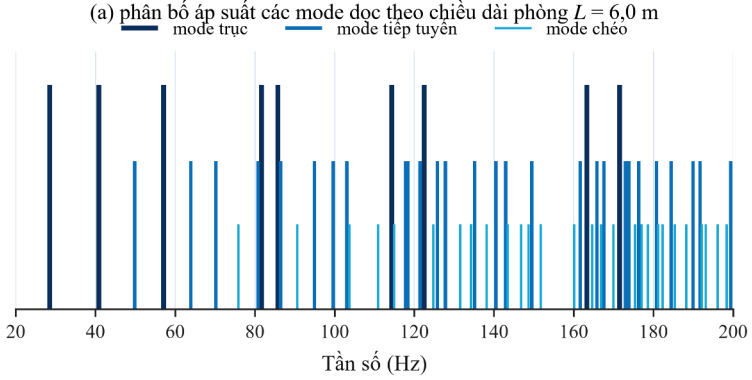
$$f_{(n_x, n_y, n_z)} = \frac{c}{2}\sqrt{\left(\frac{n_x}{L}\right)^2 + \left(\frac{n_y}{W}\right)^2 + \left(\frac{n_z}{H}\right)^2} \quad \text{(Hz)} \qquad (1.9)$$

Trong đó n_x, n_y, n_z là các số nguyên không âm (0, 1, 2, 3...). Các bộ số chỉ có một chỉ số khác không, như (1,0,0), (0,1,0), (0,0,1), được gọi là chế độ trục (Axial Modes). Chế độ trục mang năng lượng mạnh nhất (chỉ va đập giữa hai mặt phẳng đối diện) nên gây biến dạng đáp tuyến trầm nghiêm trọng nhất.

Một hệ quả thực tế cần lưu ý: tần số cộng hưởng mode trục bậc 1 theo chiều dài của một căn phòng dài 5 m rơi vào $f = 343 / (2 \times 5{,}0) = 34{,}3$ Hz (nằm trong dải Sub-bass). Tại tần số này, giữa phòng (tọa độ 2,5 m) là một nút sóng triệt tiêu áp suất. Nếu vị trí ngồi nghe được đặt đúng chính giữa phòng, người nghe sẽ thấy âm trầm ở tần số này rất yếu, dễ lầm tưởng bản thu thiếu trầm và tăng dải trầm quá mức. Khi bản phối đó được phát trên hệ thống loa khác, âm trầm sẽ bị thừa và gây ù rền.



(a) phân bố áp suất các mode dọc theo chiều dài phòng L = 6,0 m



(b) phổ mode của phòng dựng 6,0 × 4,2 × 3,0 m: các mode dồn cục ở dải trầm

Hình 1.10 Sóng đứng trong phòng kín: vị trí nút và bụng sóng, phổ các mode và giải pháp bố trí bẫy âm trầm

1.3.4 Thời gian âm vang RT60 và tần số Schroeder

Thời gian âm vang RT60 (Reverberation Time) là khoảng thời gian cần thiết, tính bằng giây, để mật độ năng lượng âm thanh trong phòng suy giảm đi 60 dB (tương đương năng lượng giảm đi một triệu lần) sau khi nguồn âm dừng phát đột ngột. Đối với các phòng có hệ số hấp thụ trung bình ở mức vừa phải ($\bar{\alpha} < 0{,}3$), RT60 được tính theo công thức Sabine:

$$RT_{60} = 0{,}161\,\frac{V}{A}\ \text{(s)}, \qquad A = \sum_i S_i\,\alpha_i \tag{1.10}$$

Trong đó: V là thể tích không gian phòng (m³); S_i là diện tích bề mặt thứ i (m²); α_i là hệ số hấp thụ âm thanh của bề mặt đó; A là tổng diện tích hấp thụ tương đương của toàn bộ căn phòng (đơn vị sabin hay m² tiêu âm). Đối với các phòng có độ hút âm rất lớn ($\bar{\alpha} > 0{,}3$ như phòng thu thoại), công thức Sabine cho sai số lớn nên cần sử dụng công thức Eyring: RT60 = 0,161 · V / [-S · ln(1 - $\bar{\alpha}$)].

Bảng 1.5 Thời gian âm vang RT60 khuyến nghị cho một số loại không gian

Loại không gian	RT60 khuyến nghị (ở dải 500 Hz – 1 kHz)	Yêu cầu âm học và Kiểm soát kỹ thuật
Phòng thu lồng tiếng / Thuyết minh (Voice booth)	0,15 s – 0,25 s	Cần không gian 'khô' (dead room), rất ít âm vang
Phòng dựng / Hòa âm nhỏ (Edit suite)	0,25 s – 0,35 s	Độ phẳng theo tần số quan trọng hơn con số tuyệt đối; kiểm soát chặt bass
Hội trường đa năng / Phòng hòa nhạc	1,00 s – 1,40 s	Cân bằng giữa độ rõ của lời nói và độ vang dày của âm nhạc
Nhà thờ cổ / Hang động vòm đá lớn	2,00 s – 6,00 s	Âm vang rất dài, làm giảm mạnh độ rõ của lời nói

VÍ DỤ 1.5 — ƯỚC LƯỢNG RT60 VÀ GIẢI PHÁP TIÊU ÂM CHO PHÒNG DỰNG PHIM

Tình huống khảo sát: một phòng dựng phim có chiều dài L = 5,0 m, chiều rộng W = 4,0 m, chiều cao H = 3,0 m. Tường và trần trát vữa sơn nhẵn (α = 0,03), sàn nhà lát gạch men (α = 0,02). Kỹ sư cần khảo sát RT60 ở 500 Hz và tính diện tích bông khoáng cần lắp đặt để đạt chuẩn RT60 = 0,3 giây.

Tính toán hiện trạng: thể tích phòng V = 5 × 4 × 3 = 60 m³.

Diện tích sàn và trần: $S_{sàn} = S_{trần} = 5 \times 4 = 20$ m².

Diện tích 4 bức tường: $S_{tường} = 2\times(5\times3) + 2\times(4\times3) = 30 + 24 = 54$ m².

Tổng diện tích hấp thụ hiện trạng: $A_{gốc} = (20 \times 0{,}02) + (20 \times 0{,}03) + (54 \times 0{,}03) = 0{,}4 + 0{,}6 + 1{,}62 = 2{,}62$ sabin.

RT60 hiện trạng = 0,161 × 60 / 2,62 ≈ 3,68 giây, nghĩa là phòng vang quá mức và không thể dùng để làm việc.

Tính toán nâng cấp đạt chuẩn: để đạt RT60 mục tiêu = 0,3 giây, tổng diện tích hấp thụ cần có là A_mục_tiêu = 0,161 × 60 / 0,3 = 32,2 sabin.

Lượng hấp thụ cần bổ sung: ΔA = 32,2 - 2,62 = 29,58 sabin.

Nếu sử dụng tấm bông khoáng Rockwool 50 mm chuyên dụng có α = 0,75 ở 500 Hz, diện tích vật liệu tiêu âm cần lắp là S_tiêu_âm = 29,58 / 0,75 ≈ 39,4 m² (tương đương phủ kín toàn bộ trần và gần 35% diện tích tường).

Bài học thực tiễn: số liệu tính toán cho thấy vài tấm mút xốp dán rải rác với mục đích trang trí không đủ để kiểm soát âm vang của một căn phòng.

Tần số Schroeder (Schroeder Frequency — f_s) là ranh giới phân định hai chế độ ứng xử âm học khác biệt của căn phòng [8]:

$$f_{\text{Schroeder}} \approx 2000\sqrt{\frac{RT_{60}}{V}} \quad \text{(Hz)} \tag{1.11}$$

Dưới tần số Schroeder ($f < f_s$), căn phòng ứng xử như một hệ dao động cộng hưởng rời rạc của các sóng đứng (Room Modes); ở vùng này, các giải pháp dùng mút xốp mỏng không có tác dụng, nên cần xử lý bằng bẫy âm trầm (Bass Trap) và bố trí vị trí ngồi nghe tối ưu. Trên tần số Schroeder ($f > f_s$), mật độ mode phòng dày đặc và giao thoa đan xen, nên căn phòng ứng xử theo quy luật âm học hình

học thống kê và được xử lý hiệu quả bằng các tấm tiêu âm bông khoáng và tán âm gỗ thông thường.

1.3.5 Tiêu chuẩn tiếng ồn nền NC/NR và giải pháp xử lý thực tế

Bên cạnh âm vang, chỉ tiêu quan trọng thứ hai của một phòng thu là tiếng ồn nền (Noise Floor), được đánh giá theo đường cong tiêu chí tiếng ồn (Noise Criteria — NC). Phòng thu tiếng nói yêu cầu mức ồn nền rất thấp, vào khoảng NC-15 đến NC-20, còn các không gian sản xuất ít khắt khe hơn có thể chấp nhận mức NC-25 đến NC-30. Bốn giải pháp xử lý âm học thường dùng gồm:

Hình 1.11 minh họa một phòng điều khiển được xử lý âm học bằng các tấm tán âm dạng thanh và các vách tường vát, phục vụ việc kiểm soát phản xạ trong phòng kiểm âm.

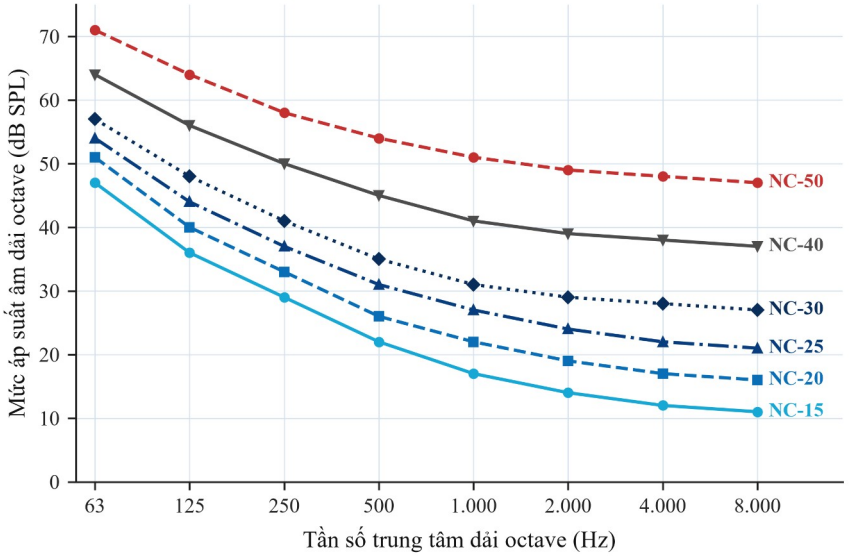


Hình 1.11 Phòng điều khiển phòng thu có tấm tán âm dạng thanh và vách vát góc

Nguồn: Iconicarecordings, CC BY-SA 4.0, Wikimedia Commons (truy cập 05/11/2026).

1. **Chăn tiêu âm (Sound Blankets): tại các không gian thu âm tạm thời có sàn đá nhẵn hoặc tường kính trơn, có thể trải thảm dạ hoặc treo chăn tiêu âm dày quanh vị trí thu để giảm dội âm vào micro**.
2. **Bố trí bẫy âm trầm góc phòng (Corner Bass Traps):** tại các góc phòng dựng (nơi cả 3 mặt phẳng tường, sàn, trần giao nhau), áp suất âm của mọi mode sóng đứng đều đạt cực đại (bụng sóng), nên việc đặt các khối bẫy trầm bằng bông khoáng tỷ trọng cao (hoặc bẫy màng cộng hưởng) tại các góc phòng đem lại hiệu quả hấp thụ năng lượng trầm cao gấp 3–4 lần so với đặt giữa tường.
3. **Quy tắc vị trí ngồi nghe tối ưu 38%:** cần tránh kê bàn dựng phim ở chính giữa phòng theo chiều dài (50% L, nút sóng mode lẻ). Vị trí tối ưu được các chuyên gia âm học quốc tế khuyến nghị là khoảng 38% chiều dài căn phòng tính từ bức tường phía trước.

4. **Tán âm và tường vát góc không song song:** các vách tường đối diện trong phòng dựng được thiết kế vát góc không song song (nghiêng từ 6° đến 12°) để triệt tiêu dội âm lặp lại (Flutter Echo). Phía sau vị trí ngồi nghe lắp đặt các tấm tán âm Schroeder (Schroeder Diffuser) bằng gỗ để phân tán đều năng lượng âm thanh, nhờ đó tạo trường âm mở rộng tự nhiên.



Hình 1.12 Hệ đường cong tiêu chí tiếng ồn nền NC và mức yêu cầu đối với từng loại không gian sản xuất

1.3.6 Nhiễu xạ, khúc xạ ngày/đêm và hiệu ứng Doppler

1. **Nhiễu xạ âm thanh (Diffraction):** là khả năng sóng âm bẻ cong và vòng qua các mép chướng ngại vật trong không gian. Mức độ nhiễu xạ tỷ lệ thuận với bước sóng: âm trầm (bước sóng vài mét) dễ dàng vòng qua các tấm vách ngăn, trong khi âm cao (bước sóng vài cm) bị chặn lại và tạo thành 'bóng âm' (acoustic shadow). Vì vậy, khi đứng sau vách tường hoặc góc phố, người nghe chỉ nhận được phần âm trầm của tiếng xe cộ.
2. **Khúc xạ do nghịch nhiệt khí quyển (Refraction):** là hiện tượng tia âm bị bẻ cong hướng lan truyền khi đi qua các tầng không khí có vận tốc âm khác nhau. Ban ngày, mặt đất bị mặt trời làm nóng nên tầng không khí sát đất nóng hơn tầng trên cao; vận tốc âm ở dưới lớn hơn ở trên khiến tia âm bị bẻ cong hướng lên trời, làm âm thanh ngoài trời nhanh chóng suy giảm. Ban đêm xảy ra hiện tượng nghịch nhiệt (nhiệt độ sát đất lạnh hơn trên cao), tia âm bị bẻ cong ngược xuống mặt đất, nên tiếng ồn từ xa vào ban đêm nghe rõ và truyền xa hơn; điều này cần lưu ý khi thu âm ngoài trời.
3. **Hiệu ứng Doppler (Doppler Effect):** là sự dịch chuyển tần số cảm nhận khi nguồn phát âm và người nghe chuyển động tương đối so với nhau với vận tốc v_s và v_o:

$$f' = f\frac{c \pm v_o}{c \mp v_s} \quad \text{(Hz)} \tag{1.12}$$

Khi xe ô tô cảnh sát hú còi tiến lại gần người nghe, tần số cảm nhận f' tăng lên (âm sắc đanh hơn); khi xe đi qua và rời xa, tần số giảm đột ngột xuống mức

trầm. Hiệu ứng này cần được tính đến khi thu âm hoặc mô phỏng âm thanh của các nguồn chuyển động.

1.4 Đơn vị decibel và các quy chuẩn đo lường

Cơ quan thính giác của con người tiếp nhận một dải áp suất âm thanh rất rộng: từ ngưỡng nghe thấy nhỏ nhất ($p_0 = 2 \cdot 10^{-5}$ Pa = 20 μPa) đến ngưỡng đau sinh lý (p = 20 Pa), chênh lệch nhau gấp 1.000.000 lần về áp suất và gấp 1.000.000.000.000 lần (một nghìn tỷ lần) về năng lượng công suất. Nếu biểu diễn dải số khổng lồ này trên thang đo tuyến tính thông thường, các bảng biểu kỹ thuật sẽ không khả thi trong vận hành. Đồng thời, cảm nhận sinh lý của tai người xấp xỉ tuân theo quy luật logarit (định luật Weber–Fechner: cảm giác tỷ lệ với logarit của đại lượng kích thích). Xuất phát từ hai lý do nền tảng đó, đơn vị decibel (dB) dựa trên logarit thập phân được áp dụng thống nhất trong toàn bộ ngành công nghiệp âm thanh.

1.4.1 Bản chất tỷ số logarit: vì sao lúc dùng hệ số 20, lúc dùng hệ số 10

Về nguyên tắc căn bản, decibel không phải là một đơn vị đo lường vật lý tuyệt đối có thứ nguyên (như mét hay kilogram), mà là một tỷ số logarit so sánh giữa đại lượng đang đo với một đại lượng quy chiếu chuẩn xác định. Vì vậy, phát biểu 'tín hiệu này là 50 dB' mà không nêu mốc quy chiếu thì không có ý nghĩa trong kỹ thuật audio.

Về bản chất vật lý, decibel được định nghĩa gốc dựa trên tỷ số công suất (Power ratio): $L = 10 \cdot \log_{10}(P / P_0)$. Tuy nhiên, công suất âm P (hoặc công suất điện của tín hiệu) luôn tỷ lệ thuận với bình phương của áp suất âm p (hoặc bình phương của điện áp V): $P \propto p^2$. Áp dụng tính chất toán học của hàm logarit: $\log_{10}(x^2) = 2 \cdot \log_{10}(x)$, thu được phép biến đổi:

$$L_p = 10\log_{10}\frac{p^2}{p_0^2} = 20\log_{10}\frac{p}{p_0} \quad \text{(dB)} \tag{1.13}$$

Quy tắc phân biệt hệ số 10 và 20 mà người đọc cần nắm vững:

1. **Dùng hệ số 20:** áp dụng hệ số 20 khi đại lượng khảo sát là đại lượng biên độ trường (áp suất âm p, điện áp micro V, dòng điện I_e, mẫu số kỹ thuật số x): $L = 20 \cdot \log_{10}(A / A_0)$.
2. **Dùng hệ số 10:** áp dụng hệ số 10 khi đại lượng khảo sát là đại lượng công suất, cường độ hay năng lượng (công suất âm P, cường độ âm I, công suất điện W, mật độ năng lượng E): $L = 10 \cdot \log_{10}(W / W_0)$.

Sự nhầm lẫn giữa hệ số 20 và 10 là sai sót tính toán phổ biến nhất của người mới học và luôn dẫn đến kết quả sai lệch gấp đôi.

1.4.2 Bảng các mốc quy chiếu decibel thông dụng

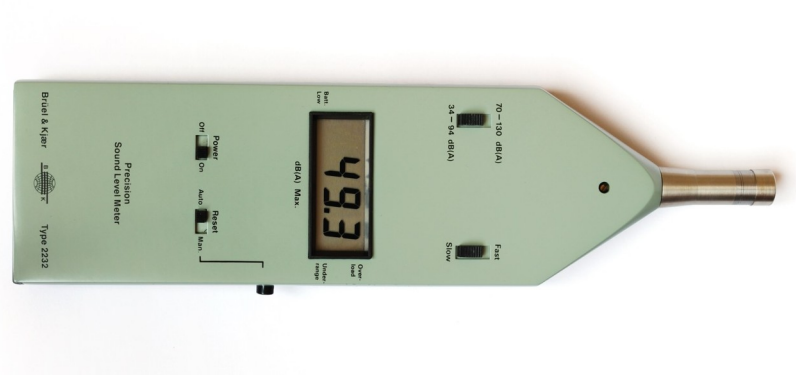
Bảng 1.6 Danh mục các mốc quy chiếu Decibel chuẩn mực trong kỹ thuật Audio

Ký hiệu chuẩn	Mốc tham chiếu quy chuẩn	Môi trường và Ngữ cảnh sử dụng chuyên nghiệp
dB SPL	p_0 = 20 µPa ($2 \cdot 10^{-5}$ Pa)	Áp suất âm thanh trong không khí; dùng cho máy đo mức âm
dBA	20 µPa + Mạch lọc trọng số A	Đo tiếng ồn môi trường, tiêu chuẩn an toàn lao động (mô phỏng 40 phon)
dBC	20 µPa + Mạch lọc trọng số C	Đo âm thanh mức lớn, hiệu chuẩn hệ thống loa kiểm âm (SMPTE ST 202: 85 dBC SPL)
dBZ	20 µPa (Đặc tuyến phẳng)	Đo kiểm âm học thuần túy tuyến tính từ 10 Hz đến 20 kHz, không bù trừ cảm nhận
dBu	V_0 = 0,7746 V RMS	Mức tín hiệu điện tương tự chuyên nghiệp (Pro Audio); +4 dBu là mức chuẩn danh định
dBV	V_0 = 1,0000 V RMS	Mức tín hiệu điện dân dụng / bán chuyên (Consumer); -10 dBV là mức chuẩn danh định
dBm	P_0 = 1 mW trên tải trở 600 Ω	Chuẩn di sản ngành viễn thông và truyền dẫn tương tự thời kỳ đầu
dBFS	Mức toàn dải số (Full Scale)	Toàn bộ hệ thống âm thanh số (DAW); 0 dBFS là trần tuyệt đối, không thể vượt qua
dBTP	Full Scale (Đỉnh thực liên mẫu)	Kiểm định True Peak chống méo rách tiếng khi nén AAC/MP3 phát hành trực tuyến
dB SIL	$I_0 = 10^{-12}$ W/m²	Mức cường độ âm thanh trong không gian (Sound Intensity Level)
dB SWL	$P_0 = 10^{-12}$ W (1 pW)	Mức công suất bức xạ âm thanh của nguồn loa (Sound Power Level)

1.4.3 Thang đo mức áp suất âm dB SPL trong thực tế

Mức áp suất âm thanh L_p (Sound Pressure Level — SPL) đo bằng dB SPL phản ánh các mức độ âm thanh quen thuộc trong thực tế [9]:

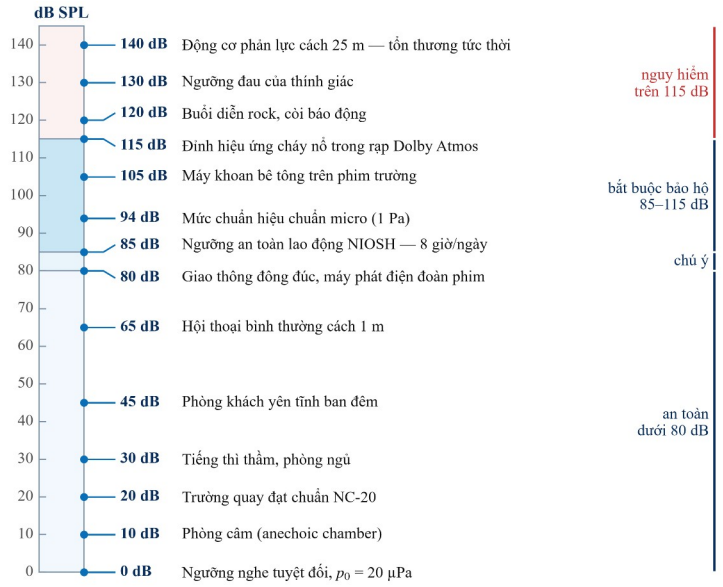
Hình 1.13 giới thiệu máy đo mức âm chính xác loại 1, thiết bị hiển thị trực tiếp mức âm theo dB(A) với micro đo ở một đầu máy.



Hình 1.13 Máy đo mức âm chính xác với micro đo ở đầu máy và màn hình hiển thị dB(A)

Nguồn: Harke, Public domain, Wikimedia Commons (đã xoay) (truy cập 05/11/2026).

1. **0 dB SPL:** ngưỡng nghe tuyệt đối của con người tại tần số 1 kHz (p = 20 µPa).
2. **20 – 30 dB SPL:** phòng thu âm, phòng cách âm tiêu chuẩn cao.
3. **40 – 45 dB SPL:** không gian nhà ở yên tĩnh, tiếng nói thì thầm.
4. **60 – 65 dB SPL:** mức âm hội thoại bình thường.
5. **75 – 85 dB SPL:** tiếng xe hơi lưu thông, tiếng ồn đường phố thông thường.
6. **85 dBC SPL:** mức hiệu chuẩn hệ thống loa kiểm âm theo tiêu chuẩn SMPTE ST 202.
7. **95 – 100 dB SPL:** xưởng cơ khí, tàu điện ngầm, tiếng người hét lớn ở cự ly gần.
8. **110 – 120 dB SPL:** concert nhạc rock, sàn nhảy, vị trí sát cụm loa sân khấu lớn; cần dùng nút tai bảo hộ thính giác.
9. **120 – 130 dB SPL:** ngưỡng đau sinh lý (Threshold of Pain); tế bào lông mao ốc tai bị tổn thương cơ học tức thì.
10. **140 – 150 dB SPL:** động cơ phản lực cất cánh cự ly gần 30 m, tiếng bom nổ; màng micro có nguy cơ rách hỏng cơ học.



Hình 1.14 Thang mức áp suất âm dB SPL từ ngưỡng nghe tuyệt đối đến ngưỡng đau, kèm giới hạn tiếp xúc nghề nghiệp

1.4.4 Công suất âm, cường độ âm và quan hệ với thanh áp

Ngoài thanh áp, còn hai đại lượng nữa mô tả độ lớn của âm thanh, và ba đại lượng này gắn chặt với nhau.

Công suất âm P là năng lượng âm đi qua một diện tích S trong một đơn vị thời gian:

$$\mathbf{P = p \cdot S \cdot v \ (W)}$$

với p là thanh áp (Pa), S là diện tích (m²), v là tốc độ dao động của phần tử môi trường tại đó (m/s).

Cường độ âm I là công suất âm đi qua một đơn vị diện tích:

$$\mathbf{I = P / S = p \cdot v \ (W/m^2)}$$

Từ hai công thức trên suy ra quan hệ ba chiều: $\mathbf{P = I \cdot S = p \cdot S \cdot v}$. Cả ba đại lượng đều biểu thị độ lớn của âm thanh; âm có năng lượng càng lớn thì công suất, cường độ và thanh áp đều càng lớn.

Bảng 1.7 Ba đại lượng công suất, cường độ, thanh áp và công thức mức tương ứng

Đại lượng	Ký hiệu	Đơn vị	Mốc tham chiếu	Công thức mức
Thanh áp	p	Pa	$p_0 = 2\times10^{-5}$ Pa	$L_p = 20\cdot\log_{10}(p/p_0)$
Cường độ âm	I	W/m²	$I_0 = 10^{-12}$ W/m²	$L_I = 10\cdot\log_{10}(I/I_0)$
Công suất âm	P	W	$P_0 = 10^{-12}$ W	$L_W = 10\cdot\log_{10}(P/P_0)$

Vì sao ba thang này cho cùng một con số ở cùng một âm

Các mốc tham chiếu được chọn có chủ ý sao cho trong trường tự do, ba thang trùng nhau. Ở ngưỡng nghe: $p = 2\times10^{-5}$ Pa cho $L_p = 0$ dB, đồng thời $I = 10^{-12}$ W/m² cho $L_I = 0$ dB. Ở ngưỡng đau: $p = 20$ Pa cho $\mathbf{L_p = 20\cdot\log_{10}(20 / 2\times10^{-5}) = 20\cdot\log_{10}(10^6) = 120\ dB}$, và $I = 1$ W/m² cho $L_I = 10\cdot\log_{10}(10^{12}) = 120$ dB. Sự trùng

khớp này khiến người làm nghề có thể nói "120 dB" mà không cần nêu rõ đang dùng thang nào.

$$I = \frac{P}{S} = \frac{p^2}{\rho c} \quad (\mathrm{W/m^2}) \tag{1.14}$$

$$L_I = 10\log_{10}\frac{I}{I_0}\ (\mathrm{dB}), \qquad I_0 = 10^{-12}\ \mathrm{W/m^2} \tag{1.15}$$

1.4.5 Phép cộng mức âm decibel

Phép cộng mức âm là phép tính dễ bị thực hiện sai trong thực hành. Vì decibel là thang đo logarit, không được cộng trực tiếp các giá trị số dB với nhau (ví dụ: 80 dB + 80 dB không bằng 160 dB). Với n nguồn âm độc lập không tương quan (uncorrelated, như hai quạt tản nhiệt, tiếng ồn phố xá), năng lượng âm thanh cộng dồn theo công suất:

$$L_\Sigma = 10\log_{10}\left(\sum_i 10^{L_i/10}\right) \quad (\mathrm{dB}) \tag{1.16}$$

Từ công thức nền tảng trên, rút ra ba quy tắc thực nghiệm quan trọng:

1. **Hai nguồn không tương quan có mức âm bằng nhau:** mức âm tổng chỉ tăng thêm đúng +3 dB: $L_{tổng} = L + 10\cdot\log_{10}(2) = L + 3{,}01$ dB (ví dụ: hai máy phát điện cùng phát 80 dB hoạt động đồng thời sẽ tạo ra mức tổng là 83 dB, không phải 160 dB). Bốn nguồn bằng nhau tăng +6 dB; mười nguồn bằng nhau tăng +10 dB.
2. **Hai tín hiệu tương quan, cùng pha hoàn toàn:** khi phát cùng một tín hiệu trên hai loa đặt sát nhau hoặc nhân đôi một track trong DAW, các dao động áp suất cộng trực tiếp trước khi bình phương công suất: $L_{tổng} = L + 20\cdot\log_{10}(2) = L + 6{,}02$ dB (tăng đúng +6 dB).
3. **Quy tắc chênh lệch 10 dB:** nếu hai nguồn âm chênh lệch nhau từ 10 dB trở lên, nguồn âm yếu hơn chỉ đóng góp tối đa 0,4 dB vào mức tổng và có thể bỏ qua trong tính toán thực tế (đây là cơ sở định lượng của quy tắc 3:1).

VÍ DỤ 1.2 — PHÉP CỘNG TIẾNG ỒN VÀ CHIẾN LƯỢC XỬ LÝ TIẾNG ỒN PHÒNG THU

Bối cảnh hiện trường: trong một phòng thu âm, kết quả đo ghi nhận máy điều hòa nhiệt độ gây tiếng ồn 62 dB SPL, quạt tản nhiệt của thiết bị gây 58 dB SPL và tiếng ồn giao thông ngoài đường lọt qua cửa đạt 55 dB SPL. Yêu cầu đặt ra là mức ồn nền tổng thể phải dưới 35 dB SPL; bài toán là tính mức ồn tổng hiện tại và đề xuất hướng xử lý.

Tính toán mức ồn tổng: ba nguồn không tương quan nên được cộng theo công suất:

10^(62/10) = 1.584.893

10^(58/10) = 630.957

10^(55/10) = 316.228

Tổng năng lượng = 2.532.078

$L_{tổng} = 10 \cdot \log_{10}(2.532.078) \approx 64{,}0$ dB SPL.

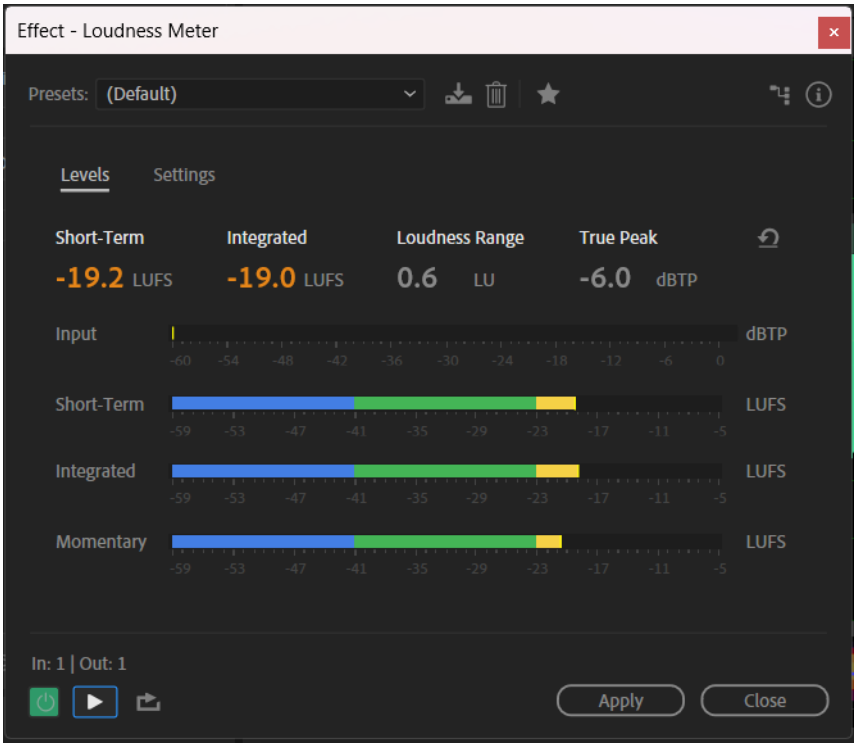
Nhận xét và phân tích: mức ồn tổng 64 dB chỉ cao hơn nguồn ồn lớn nhất (máy điều hòa 62 dB) đúng 2 dB, vì nguồn ồn lớn nhất luôn chi phối trường âm.

Chiến lược can thiệp chuẩn mực: để giảm mức ồn từ 64 dB về 35 dB (giảm tới 29 dB), cần xử lý nguồn mạnh nhất trước tiên. Việc tắt máy điều hòa trong lúc thu đưa mức ồn xuống còn 59,7 dB; sau đó thay quạt loại êm hơn và làm kín cửa cách âm. Bài học rút ra là phải xử lý nguồn ồn lớn nhất trước khi đầu tư các biện pháp khác như tiêu âm tường.

1.4.6 Tiêu chuẩn đo độ to hiện đại LUFS/LKFS theo ITU-R BS.1770

Đo lường bằng dB SPL hay dBFS là các thước đo vật lý kỹ thuật thuần túy, không phản ánh được cảm giác thực tế của người nghe về độ to (Loudness). Trong kỷ nguyên phát thanh truyền hình số và các nền tảng phân phối trực tuyến (Streaming), ngành công nghiệp âm thanh toàn cầu đã thống nhất áp dụng chuẩn đo độ to chuẩn hóa quốc tế: LUFS (Loudness Units relative to Full Scale, chuẩn Châu Âu EBU) hoặc LKFS (Loudness K-weighted Full Scale, chuẩn Bắc Mỹ ATSC). Về mặt toán học, 1 LUFS hoàn toàn đồng nhất với 1 LKFS.

Hình 1.15 cho thấy đồng hồ Loudness Meter của Adobe Audition sau khi đo một đoạn thoại tổng hợp dài sáu giây: độ to tích hợp đạt −19,0 LUFS, dải độ to 0,6 LU và đỉnh thực −6,0 dBTP, cùng các thanh ngắn hạn, tích hợp và tức thời.



Hình 1.15 Đồng hồ đo độ to LUFS (Loudness Meter) trong Adobe Audition; nguồn âm là tệp thoại mẫu tổng hợp

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Audition 2026, bản 26.5), 07/11/2026.

Thuật toán đo độ to tiêu chuẩn ITU-R BS.1770 (phiên bản BS.1770-4/-5) [1] bao gồm bốn giai đoạn xử lý cốt lõi:

1. **Mạch lọc trọng số K (K-weighting):** tín hiệu âm thanh đi qua hai tầng lọc nối tiếp, gồm một bộ lọc High-Shelf tần số cao (+4 dB từ khoảng 2 kHz trở lên) nhằm mô phỏng hiệu ứng tán xạ làm tăng thanh áp của đầu và vành tai người, và một bộ lọc High-Pass cắt bỏ năng lượng dưới 100 Hz nhằm mô phỏng đặc tính kém nhạy của thính giác với dải trầm.
2. **Tính bình phương trung bình và trọng số kênh:** năng lượng các kênh L, R, C được nhân hệ số trọng số 1,0, các kênh vòm (Surround) được nhân trọng số 1,41 (+1,5 dB) do tai người nhạy cảm hơn với âm thanh đến từ phía sau, còn kênh siêu trầm LFE bị loại hoàn toàn khỏi phép tính độ to.
3. **Thuật toán phân vùng cổng (Gating Mechanism):** cơ chế đóng cổng hai cấp độ loại trừ các đoạn khoảng lặng (như quãng ngắt nghỉ giữa các câu thoại hay đoạn fade out), trong đó cổng tuyệt đối (Absolute Gate) loại bỏ mọi khối tín hiệu dưới -70 LUFS và cổng tương đối (Relative Gate) tiếp tục loại bỏ các khối tín hiệu thấp hơn 10 LU so với mức trung bình sơ bộ. Nhờ cơ chế Gating, một chương trình có nhiều đoạn tĩnh lặng không bị đánh giá sai lệch là quá nhỏ.
4. **Tích hợp theo ba khung thời gian:** phép đo được thực hiện trên ba khung thời gian là Momentary (đo tức thời trong 400 ms), Short-term (đo trượt trong 3 giây) và Integrated (đo tích hợp liên tục từ đầu đến cuối toàn bộ thời lượng tác phẩm). Chỉ số Integrated LUFS là chỉ số được các tiêu chuẩn phát hành dùng làm mức đích.

Hai chỉ số kỹ thuật cần được theo dõi song song trên đồng hồ đo độ to (Loudness Meter):

1. **Loudness Range (LRA):** dải độ to đo bằng đơn vị LU, mô tả mức chênh lệch năng lượng giữa đoạn tĩnh lặng nhất và đoạn bùng nổ nhất của tác phẩm. Nội dung nghe trên thiết bị di động nên giữ LRA hẹp để người nghe không phải liên tục tăng giảm âm lượng.
2. **Mức đỉnh thực (True Peak, dBTP):** khác với Sample Peak (chỉ đọc giá trị các mẫu rời rạc), True Peak sử dụng thuật toán nội suy nâng tần số lấy mẫu (Oversampling) để tái tạo chính xác các đỉnh sóng hình sin liên tục xuất hiện sau mạch giải mã D/A. Một tệp âm thanh có Sample Peak nằm an toàn ở 0 dBFS vẫn có thể vọt lên +1,5 dBTP khi nén sang định dạng MP3 hoặc AAC, gây méo vỡ tiếng trên thiết bị của người nghe. Do đó, mọi tiêu chuẩn phát hành đều quy định mức trần True Peak tối đa từ -1,0 dBTP đến -2,0 dBTP.

Bảng 1.8 Tiêu chuẩn đo độ to LUFS và Trần True Peak trên các nền tảng phát hành

Nền tảng / Chuẩn phát hành	Mức Integrated đích	Trần True Peak Max	Quy chuẩn kỹ thuật chuyên môn
EBU R128 (Phát thanh – TV Châu Âu) [2]	-23,0 LUFS (±0,5 LU)	-1,0 dBTP	Tiêu chuẩn pháp lý truyền hình Châu Âu, đo toàn bộ chương trình
ATSC A/85 (Truyền hình Mỹ) [4]	-24,0 LKFS (±1,0 LU)	-2,0 dBTP	Tuân thủ Đạo luật CALM Act kiểm soát âm lượng quảng cáo thương mại
Netflix	-27,0 LKFS (Dialog-gated)	-2,0 dBTP	Đo độ to dựa trên phần lời thoại (Dialog-gated)
YouTube	-14,0 LUFS	-1,0 dBTP	Tự động hạ âm lượng nếu to hơn -14 LUFS; không tự tăng nếu nhỏ hơn
Spotify	-14,0 LUFS	-1,0 dBTP	Tính năng Loudness Normalization bật mặc định ở mức chuẩn
Apple Music (Stereo)	-16,0 LUFS	-1,0 dBTP	Tiêu chuẩn Apple Digital Masters (ADM) bảo toàn độ động
Apple Music (Spatial / Atmos)	-18,0 LUFS	-1,0 dBTP	Tiêu chuẩn bắt buộc cho toàn bộ bản phối âm thanh vòm Dolby Atmos
TikTok / Facebook Reels / Shorts	-14,0 LUFS	-1,0 dBTP	Nội dung ngắn, tối ưu hóa phát trên loa ngoài của điện thoại thông minh
AES TD1004 (Podcast chuyên nghiệp) [5]	-16,0 LUFS (Stereo)	-1,0 dBTP	Khuyến nghị của Hội Kỹ sư Âm thanh Quốc tế (-19 LUFS cho mono)

Bản chất cuộc chiến độ to (Loudness War) và tư duy mastering hiện đại

Bối cảnh lịch sử: trước kỷ nguyên phát hành trực tuyến, trên sóng radio và đĩa CD, bản nhạc có âm lượng lớn hơn sẽ lập tức gây ấn tượng mạnh hơn với người nghe. Điều này dẫn tới cuộc 'chiến tranh độ to' kéo dài suốt thập niên 1990–2000, khi các kỹ sư liên tục ép nén (Brickwall Limiting) để đẩy RMS lên sát trần và bóp Crest Factor xuống chỉ còn 5–6 dB.

Sự sụp đổ của tư duy làm to trong kỷ nguyên Streaming: ngày nay, các nền tảng hàng đầu (YouTube, Spotify, Apple Music, Netflix) đều tích hợp thuật toán chuẩn hóa độ to tự động (Loudness Normalization). Nếu người làm âm thanh master một bản nhạc ở mức rất lớn là -8 LUFS rồi tải lên YouTube (-14 LUFS), thuật toán của YouTube sẽ tự động kéo giảm toàn bộ bản mix xuống -6 dB để đưa về đúng -14 LUFS.

Hậu quả: bản mix bị ép nén quá mức, khi bị kéo nhỏ lại, nghe bẹp, thiếu lực, mờ đục và mất không gian sống động; trong khi đó một bản master tinh tế giữ được dải động ở -14 LUFS sẽ khi phát lại nghe thoáng, tiếng trống nảy và có uy lực hơn nhiều. Mục tiêu của kỹ thuật âm thanh hiện đại là tối ưu chất lượng dải động và độ rõ nét, không phải chạy đua con số LUFS lớn nhất.

1.4.7 Định luật bình phương nghịch đảo và khoảng cách đặt micro

Khi sóng âm lan truyền tự do ngoài trời từ một nguồn điểm, năng lượng bức xạ phân bổ đều trên diện tích mặt cầu $S = 4\pi r^2$. Theo Định luật bình phương nghịch đảo (Inverse Square Law), mật độ cường độ năng lượng tỷ lệ nghịch với bình phương khoảng cách ($I \propto 1/r^2$). Chuyển sang thang đo Decibel thanh áp:

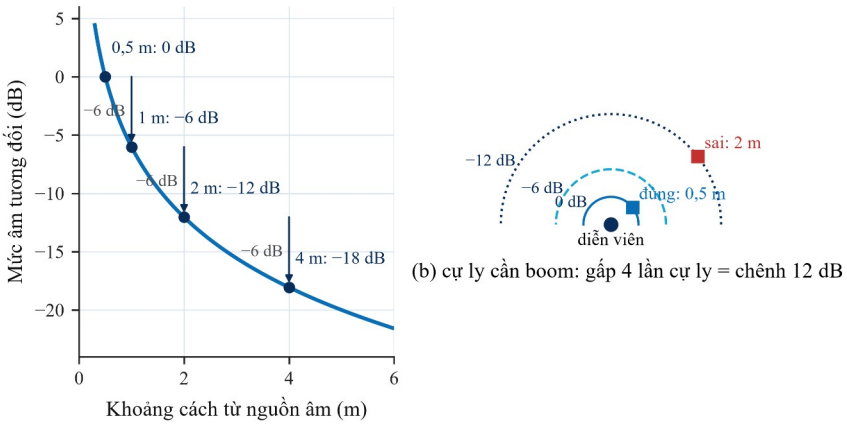
$$\Delta L = 20\log_{10}\frac{r_1}{r_2} \quad \text{(dB)} \tag{1.17}$$

Hai mốc suy hao kinh điển mà người đọc cần nắm vững là:

1. **Quy tắc nhân đôi cự ly (-6 dB):** mỗi khi khoảng cách từ nguồn âm đến micro tăng gấp đôi (từ r lên 2r), mức áp suất âm thanh thu được suy giảm đúng 6,02 dB (thường làm tròn thành -6 dB): $\Delta L = 20 \cdot \log_{10}(1/2) \approx -6{,}02$ dB.
2. **Quy tắc tăng 10 lần cự ly (-20 dB):** mỗi khi khoảng cách tăng gấp 10 lần (từ r lên 10r), mức áp suất âm thanh thu được suy giảm đúng 20 dB: $\Delta L = 20 \cdot \log_{10}(1/10) = -20$ dB.

Bảng 1.9 Bảng tra nhanh mức suy hao âm thanh theo khoảng cách

Khoảng cách so với mốc ban đầu (r)	Mức áp suất âm thay đổi (ΔL)
Cự ly ban đầu r = 0,5 m	0 dB (Mốc chuẩn)
Tăng gấp 2 lần (r = 1,0 m)	-6 dB
Tăng gấp 4 lần (r = 2,0 m)	-12 dB
Tăng gấp 10 lần (r = 5,0 m)	-20 dB
Tăng gấp 40 lần (r = 20,0 m)	-32 dB
Thu hẹp còn 1/2 cự ly (r = 0,25 m)	+6 dB



(a) mỗi lần nhân đôi cự ly, mức âm giảm 6 dB

(b) cự ly cần boom: gấp 4 lần cự ly = chênh 12 dB

Hình 1.16 Định luật bình phương nghịch đảo: mức âm giảm 6 dB mỗi lần nhân đôi cự ly, và hệ quả với cự ly đặt micro

1.4.8 Dải động, sàn nhiễu và kỹ thuật gain staging

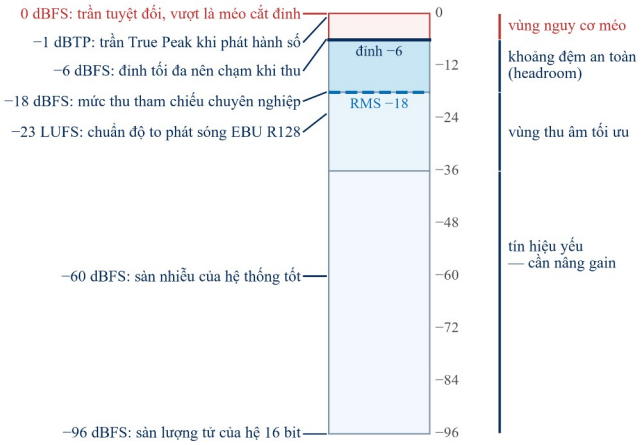
Một chuỗi thiết bị âm thanh từ micro, mixer đến máy ghi âm luôn có bốn giới hạn vật lý bất biến [11], [12]:

1. **Sàn nhiễu (Noise Floor):** mức tín hiệu nhiễu điện tử cố hữu phát sinh do chuyển động nhiệt của các hạt electron trong mạch điện trở và bán dẫn khi không có tín hiệu âm thanh hữu ích.
2. **Dải động (Dynamic Range):** khoảng cách tính bằng decibel từ sàn nhiễu lên tới mức tín hiệu cực đại trước khi xuất hiện hiện tượng méo cắt đỉnh (clipping). Dải động lý thuyết của hệ thống 16-bit (CD) là 96 dB, còn hệ thống 24-bit chuyên nghiệp đạt 144 dB (thực tế khoảng 120 dB do giới hạn nhiệt analog).
3. **Tỷ số tín hiệu trên tạp âm (SNR):** tỷ số giữa mức tín hiệu làm việc danh định và sàn nhiễu.
4. **Khoảng dự trữ an toàn (Headroom):** khoảng dự trữ tính bằng dB từ mức tín hiệu làm việc danh định lên tới trần méo 0 dBFS.

Quy tắc phân bổ khuếch đại (Gain Staging) chuẩn mực: trong toàn bộ quy trình ghi âm số 24-bit hiện đại, mức tín hiệu danh định của tiếng nói nên được căn chỉnh sao cho giá trị hiệu dụng dao động quanh mức -18 dBFS RMS (chuẩn Châu Âu) hoặc -20 dBFS RMS (chuẩn SMPTE Mỹ). Cách đặt Gain này chừa lại đúng 18–20 dB Headroom an toàn cho các đỉnh tức thời khi người nói bất ngờ cười lớn hoặc nói to mà không bị méo cắt đỉnh clipping, đồng thời vẫn giữ tín hiệu cách xa sàn nhiễu tới hơn 80 dB.

Đặc điểm của máy ghi âm 32-bit Float thế hệ mới: các dòng máy ghi âm hiện trường thế hệ mới (như Sound Devices MixPre II, Zoom F6, Zoom F8n Pro) sử dụng định dạng số 32-bit dấu phẩy động (32-bit Float) có dải động lý thuyết lên tới hơn 1.500 dB, lớn đến mức tín hiệu về mặt toán học không bao giờ bị clip trong miền số. Tuy nhiên, màng rung cơ học của micro và tầng mạch tiền khuếch đại tương tự (analog preamp) phía trước bộ biến đổi ADC vẫn có giới hạn chịu đựng vật lý thực tế: nếu thanh áp vượt quá ngưỡng chịu đựng cơ học của micro (khoảng

130 dB SPL), màng micro sẽ bị méo dạng phi tuyến trước khi tín hiệu kịp chuyển sang dạng số. Do đó, nguyên tắc phân bổ Gain staging và chừa khoảng dự trữ headroom vẫn giữ nguyên giá trị nghề nghiệp.



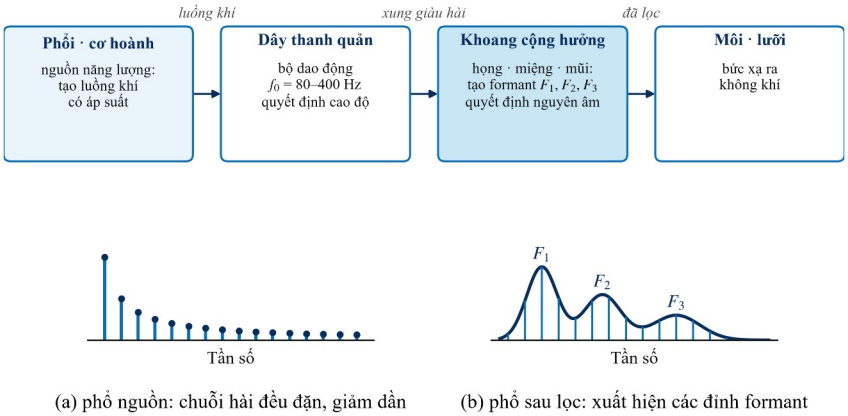
Hình 1.17 Thang decibel toàn dải dBFS, khoảng đệm an toàn và các mốc chuẩn trong quy trình gain staging

1.5 Cơ chế phát âm giọng nói và các dải tần cộng hưởng (Formants)

Dưới góc độ âm học vật lý, giọng nói con người hoạt động theo Mô hình nguồn – bộ lọc (Source-Filter Model): phổi và dây thanh quản đóng vai trò nguồn tạo dao động kích thích ban đầu giàu hài âm, trong khi toàn bộ đường dẫn khí phía trên (hầu họng, khoang miệng, khoang mũi) đóng vai trò bộ lọc cộng hưởng định hình âm sắc.

1.5.1 Cấu tạo cơ quan phát âm và hoạt động của dây thanh quản

Luồng không khí từ phổi được đẩy qua khí quản dưới áp lực của cơ hoành và lồng ngực. Tại thanh quản, không khí đi qua khe thanh môn nằm giữa hai nếp cơ gọi là Dây thanh quản (Dây thanh đới, Cordes vocales). Dây thanh quản không dao động tự do như dây đàn vĩ cầm mà liên tục khép lại rồi bật mở nhịp nhàng dưới tác động của áp lực khí động học và hiệu ứng Bernoulli. Quá trình đóng mở liên hoàn này băm luồng khí nén thành từng đợt xung áp suất phụt ra liên tiếp với tần số cơ bản f_0 dao động từ 80 Hz đến 1.100 Hz, tùy thuộc vào độ dày, độ căng của cơ thanh quản và giới tính của người nói.

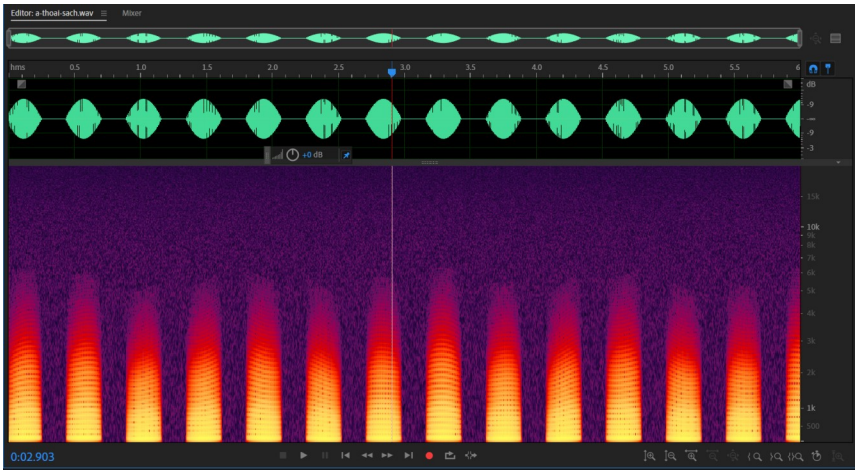


Hình 1.18 Cơ quan phát âm của người theo mô hình nguồn – bộ lọc: dây thanh quyết định cao độ, khoang cộng hưởng quyết định nguyên âm

1.5.2 Hộp cộng hưởng khoang miệng, khoang mũi và khái niệm formant

Xung áp suất thô sau khi đi qua thanh quản tiếp tục truyền vào một hệ thống các khoang rỗng gồm hầu, khoang miệng, vòm khẩu cái và hốc mũi. Hệ thống này hoạt động như một chuỗi Hộp cộng hưởng âm học (Acoustic Resonators), khuếch đại mạnh một số dải tần số nhất định và triệt tiêu các dải tần số khác. Các đỉnh cộng hưởng khuếch đại này được gọi là Formants (dải tần số tạo âm), ký hiệu tuần tự là F1, F2, F3, F4...

Hình 1.19 là phổ đồ của một chuỗi âm tiết mẫu trong chế độ Spectral Frequency Display của Adobe Audition: mỗi âm tiết hiện thành một cụm năng lượng, tập trung ở dải thấp và giảm dần về phía cao, đúng với hình ảnh các dải cộng hưởng (formant) của tiếng nói.



Hình 1.19 Phổ đồ của đoạn thoại mẫu trong Adobe Audition; nguồn âm là tệp thoại tổng hợp, trục đứng là tần số

Nguồn: ảnh chụp màn hình do người soạn thực hiện (Adobe Audition 2026, bản 26.5), 07/11/2026.

Cặp formant F1 và F2 đóng vai trò quyết định trong việc nhận diện nguyên âm: nguyên âm /i/ có F1 rất thấp (~250 Hz) và F2 rất cao (~2.300 Hz), trong khi nguyên âm /a/ có F1 cao (~700 Hz) và F2 ở mức trung bình (~1.100 Hz). Vị trí của các formant phụ thuộc vào cấu trúc giải phẫu và hình dạng khẩu hình (vị trí hàm, lưỡi, môi), gần như độc lập với cao độ f₀. Vì vậy, khi sử dụng các phần mềm biến đổi cao độ (Pitch Shifting) để xử lý giọng nói, nếu không kích hoạt chế độ bảo toàn

formant (Formant-preserving mode), giọng nói khi nâng tông sẽ biến thành 'giọng sóc chuột Chipmunk' còn khi hạ tông sẽ thành 'giọng quái vật khổng lồ', do các đỉnh formant bị dịch chuyển phi tự nhiên.

1.5.3 Phân loại dải tần các chất giọng

Bảng 1.10 Phân loại các chất giọng và đặc tính xử lý âm thanh

Chất giọng (Voix)	Dải tần cơ bản f_0	Formants đặc trưng
Basse (Nam trầm)	80 Hz – 120 Hz	Cộng hưởng mạnh 2,4 kHz
Baryton (Nam trung)	100 Hz – 180 Hz	Cộng hưởng mạnh 2,6 kHz
Ténor (Nam cao)	140 Hz – 280 Hz	Cộng hưởng mạnh 2,8 kHz
Alto / Mezzo (Nữ trung)	160 Hz – 300 Hz	Cộng hưởng mạnh 2,9 kHz
Soprano (Nữ cao)	200 Hz – 350 Hz	Cộng hưởng mạnh 3,2 kHz
Giọng trẻ em	250 Hz – 400 Hz	Cộng hưởng 3,5 kHz+
Vùng rõ lời thoại (Intelligibility)	1.000 Hz – 4.000 Hz	Đỉnh F2 và F3
Âm xuýt (Sibilance: s, x, ch)	5.000 Hz – 9.000 Hz	Gió ma sát răng

1.5.4 Kỹ thuật khoét tần số và kiểm soát hiệu ứng cận âm

Khi nhạc nền phát song song với lời thoại, việc chỉ đẩy âm lượng track thoại lên cao sẽ khiến bản phối trở nên ồn và chói tai. Một giải pháp thường dùng là kỹ thuật khoét tần số (Frequency Carving EQ): dùng bộ lọc tham số (Parametric EQ) để cắt nhẹ từ -2 dB đến -4 dB, với độ rộng băng thông vừa phải ($Q \approx 1,5$), trên track nhạc nền tại dải formant nhận diện giọng nói (vùng 1 kHz đến 3,5 kHz). Thao tác này tạo ra một khoảng trống trên phổ tần số, nhờ đó lời thoại nổi lên rõ nét mà không cần tăng tổng mức âm lượng.

Hiệu ứng cận âm (Proximity Effect) và âm nổ (Plosives) khi thu tiếng nói:

1. **Hiệu ứng cận âm (Proximity Effect):** xuất hiện trên các dòng micro có tính định hướng (Cardioid, Supercardioid, Shotgun). Khi micro đặt ở cự ly rất gần miệng người nói (dưới 15–20 cm), đáp tuyến tần số vùng trầm (dưới 150 Hz) bị tăng mạnh từ +6 dB đến +15 dB. Hiệu ứng này có thể được khai thác để tạo chất giọng phát thanh viên ấm và dày, nhưng cũng là nguyên nhân khiến giọng nói bị ùng khi người nói tiến quá sát micro.
2. **Âm nổ (Plosives):** khi người nói phát âm các phụ âm bật hơi mạnh như /p/, /b/, /t/, một luồng khí nén tốc độ cao đập trực tiếp vào màng micro, tạo ra xung áp suất tần số rất thấp làm màng rung mạnh và gây tiếng 'bụp' vỡ tiếng. Biện pháp khắc phục là luôn sử dụng màng lọc gió (Pop Filter) hoặc chụp mút lọc âm, đồng thời hướng trục micro lệch góc khoảng 15° đến 30° so với hướng luồng hơi từ miệng người nói.

1.6 Sinh lý thính giác và tâm lý âm học

Âm thanh xét đến cùng là cảm giác thính giác chủ quan được tái dựng trong vỏ não của người nghe. Mọi quyết định, từ việc chọn vị trí đặt micro, xử lý âm sắc

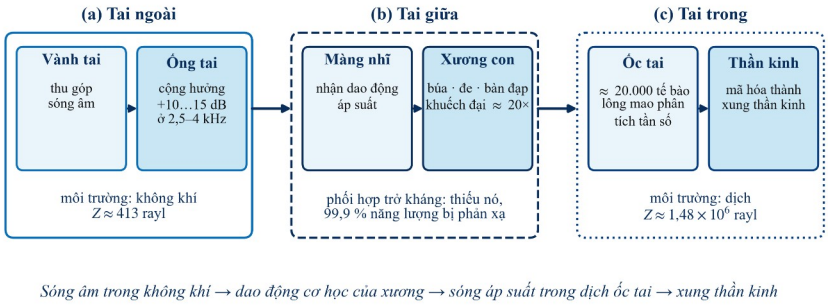
cho đến thiết lập mức âm lượng khi hòa âm, đều cần được xem xét dưới góc độ cấu tạo giải phẫu của tai người và các quy luật tâm sinh lý âm học.

1.6.1 Cấu tạo giải phẫu ba khoang tai và cơ chế phối hợp trở kháng

Cơ quan thính giác của con người gồm ba khu vực liên hoàn, mỗi khu vực có chức năng cơ sinh học tinh vi:

1. **Tai ngoài (Outer Ear):** gồm vành tai (Pinna), có cấu trúc sụn với các nếp gấp bất đối xứng giúp thu sóng áp suất và tạo ra các biến đổi phổ tần số đặc thù (cơ sở để não định vị hướng trên/dưới và trước/sau), nối tiếp vào ống tai ngoài dài khoảng 2,5 cm, hoạt động như một ống cộng hưởng hở một đầu. Ống tai tạo nên một đỉnh cộng hưởng tự nhiên, khuếch đại thanh áp từ +10 dB đến +15 dB tại dải tần số 2,5 kHz đến 4,0 kHz, và đây là nguyên nhân trực tiếp khiến tai người nhạy cảm nhất ở dải tần số này.
2. **Tai giữa (Middle Ear):** gồm màng nhĩ (Tympanic Membrane) mỏng, tiếp nhận biến thiên áp suất không khí và chuyển thành dao động cơ học. Nối tiếp sau màng nhĩ là chuỗi ba xương nhỏ nhất của cơ thể người: xương búa (Malleus), xương đe (Incus) và xương bàn đạp (Stapes). Chuỗi xương con hoạt động như một hệ đòn bẩy cơ học kép, kết hợp với sự chênh lệch diện tích giữa màng nhĩ (khoảng 55 mm^2) và đế xương bàn đạp (khoảng 3,2 mm^2), tạo nên hệ số khuếch đại áp lực hơn 20 lần. Chức năng thiết yếu của tai giữa là phối hợp trở kháng âm học (Impedance Matching) giữa môi trường không khí loãng của ống tai ($Z \approx 413$ rayl) và môi trường dịch lỏng nhớt trong ốc tai ($Z \approx 1.500.000$ rayl); nếu không có chuỗi xương đòn bẩy này, hơn 99,9% năng lượng sóng âm sẽ bị phản xạ trở lại và con người gần như điếc hoàn toàn.
3. **Tai trong (Inner Ear):** là một mê cung xoắn hình vỏ ốc (Cochlea) chứa đầy nội dịch và màng đáy (Basilar Membrane). Trên màng đáy có khoảng 20.000 tế bào lông mao (Hair Cells) nối trực tiếp với dây thần kinh thính giác số VIII. Ốc tai hoạt động như một bộ phân tích phổ cơ học theo vị trí (Tonotopic Organization): vùng đáy ốc tai (dày và căng) cộng hưởng với các tần số cao (10–20 kHz), trong khi vùng đỉnh ốc tai (mềm và rộng) cộng hưởng với các tần số trầm (20–200 Hz). Chuyển động uốn lượn của chất dịch kích thích tế bào lông mở các kênh ion, chuyển dao động cơ học thành các xung điện thần kinh truyền thẳng lên thùy thái dương của não bộ.

> **Cảnh báo nghề nghiệp: tế bào lông mao thính giác không có khả năng tái sinh.**
> Về sinh lý học, tế bào lông mao trong ốc tai người là tế bào thần kinh chuyên biệt; một khi đã gãy rụng hoặc tổn thương cơ học do tiếp xúc với âm thanh cường độ lớn, chúng hoàn toàn không có khả năng tự tái tạo hay phục hồi.
>
> Hậu quả tích lũy là tổn thương thính giác nghề nghiệp (Noise-Induced Hearing Loss, NIHL) diễn ra âm thầm, bắt đầu bằng triệu chứng ù tai (tinnitus) sau giờ làm việc và mất dần thính lực ở dải tần số 4.000 Hz. Đối với người làm âm thanh, đôi tai là công cụ làm việc không thể thay thế, nên việc bảo vệ thính giác phải được xem là kỷ luật nghề nghiệp bắt buộc.



Hình 1.20 Cấu tạo ba khoang tai, đường đi của tín hiệu và cơ chế phối hợp trở kháng âm học

1.6.2 Đường cong đẳng âm ISO 226 và mức âm lượng kiểm âm

Cảm giác về độ to (Loudness) của tai người phi tuyến tính theo tần số. Công trình lịch sử của Fletcher và Munson (1933) cùng tiêu chuẩn quốc tế hiện hành ISO 226:2023 [3] đã xác lập hệ thống các đường cong đẳng âm (Equal-Loudness Contours). Mỗi đường cong biểu diễn mức áp suất âm thanh (dB SPL) cần có ở từng tần số để tạo cùng một cảm giác độ to như ở tần số quy chiếu chuẩn 1.000 Hz (đo bằng đơn vị Phon):

1. **Đơn vị Phon:** một âm thanh có độ to n phon nghe to tương đương một sóng sin 1.000 Hz phát ở mức n dB SPL; chẳng hạn, một âm 1 kHz ở mức 60 dB SPL có độ to là 60 phon. Tuy nhiên, nếu tần số giảm xuống 100 Hz, để tai nghe cùng độ to 60 phon thì mức thanh áp thực tế phải tăng lên 68–70 dB SPL (tăng thêm gần 10 dB). Ở tần số siêu trầm 30 Hz, mức thanh áp thực tế phải tăng tới 85 dB SPL mới nghe to bằng 60 dB ở 1 kHz.
2. **Đơn vị Sone:** là thang đo tỷ lệ cảm nhận trực tiếp, trong đó 1 sone được định nghĩa bằng độ to của âm 40 phon. Cứ mỗi lần độ to tăng thêm 10 phon thì số sone tăng gấp đôi (2 sone = 50 phon, 4 sone = 60 phon, 8 sone = 70 phon). Điều này định lượng hóa một nguyên tắc quan trọng: phải tăng mức âm lên khoảng +10 dB thì người nghe mới cảm nhận âm thanh 'to gấp đôi'.

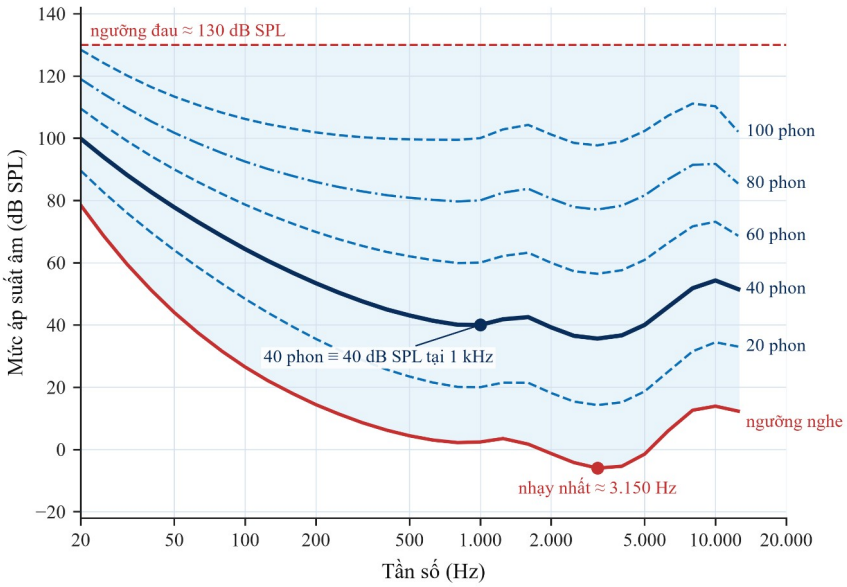
Đáp tuyến của tai thay đổi theo mức âm lượng: khi nghe ở mức nhỏ (dưới 60 dB SPL), các đường cong đẳng âm rất dốc và tai kém nhạy với dải trầm lẫn dải cao, còn khi nghe ở mức lớn (trên 85 dB SPL), các đường cong đẳng âm phẳng hơn nhiều. Hiện tượng sinh lý này là lý do cần cố định mức âm lượng kiểm âm khi hòa âm:

Mức âm lượng kiểm âm chuẩn (85 dBC SPL / 79 dBC SPL)

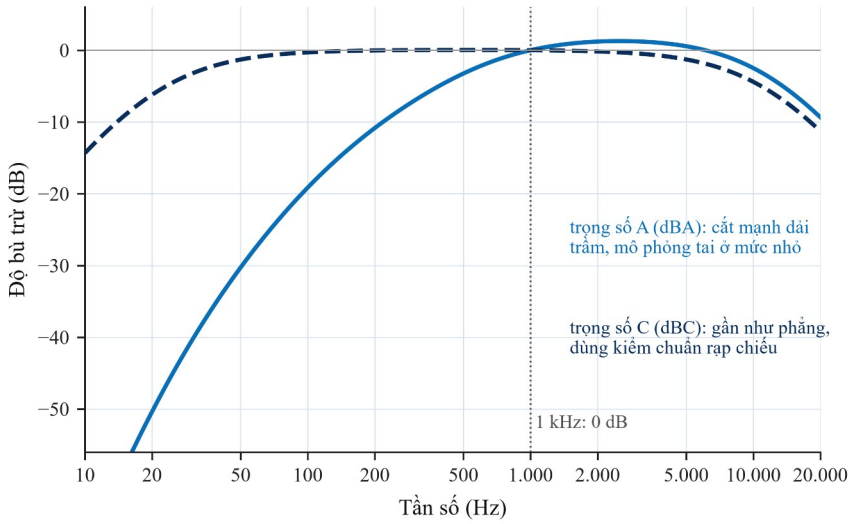
Tiêu chuẩn SMPTE ST 202 quy định dùng máy đo mức âm đặt ở trọng số C (C-weighted, Slow response) để hiệu chuẩn mức phát của hệ thống loa kiểm âm tại vị trí nghe trung tâm: 85 dBC SPL cho mỗi kênh chính (Left, Center, Right) khi phát tín hiệu nhiễu hồng danh định (Pink Noise ở mức -20 dBFS RMS). Đối với các phòng dựng nhỏ (dưới 150 m^3), mức hiệu chuẩn thường dùng là 79 dBC SPL [6].

Rủi ro nghề nghiệp khi kiểm âm sai âm lượng: nếu hòa âm ở mức quá nhỏ (ví dụ 65 dB), do đường cong đẳng âm lõm ở dải trầm, tai sẽ nhận thấy bản mix thiếu bass/treble và người hòa âm có xu hướng tăng thừa EQ ở hai đầu dải; khi phát lại ở mức lớn, âm thanh sẽ gắt, chói và thừa trầm. Ngược lại, nếu hòa âm liên tục ở mức quá to, người hòa âm sẽ gọt bớt bass, và khi phát lại ở mức nhỏ bản mix sẽ mỏng và nghèo âm sắc.

Quy tắc an toàn: nên cố định mức volume của phòng dựng ở mức chuẩn (85 dBC hoặc 79 dBC) và chỉ kiểm tra nhanh bản mix ở mức nhỏ để đánh giá độ cân bằng của lời thoại.



Hình 1.21 Hệ đường cong đẳng âm dựng theo bảng tham số chuẩn ISO 226 và vùng nghe của tai người



Hình 1.22 Đặc tuyến mạch lọc trọng số dBA (cắt trầm mạnh) đối chiếu dBC (gần như phẳng)

1.6.3 Các ngưỡng chênh lệch cảm nhận của thính giác

Việc nắm vững khả năng phân biệt mức chênh lệch âm lượng của tai người là cơ sở cho mọi thao tác đẩy fader trong hòa âm:

Bảng 1.11 Các ngưỡng chênh lệch cảm nhận của thính giác và Ý nghĩa nghiệp vụ

Mức chênh lệch âm lượng	Cảm nhận sinh lý của tai người	Ý nghĩa thực tiễn trong hòa âm
Dưới 1 dB	Hoàn toàn không thể phát hiện trong điều kiện nghe bình thường	Các thao tác chỉnh fader hay volume dưới 1 dB là vô nghĩa trong bản phối
Khoảng 1 dB (JND)	Ngưỡng phát hiện tối thiểu (Just Noticeable Difference)	Chỉ nhận ra khi làm bài test chuyển đổi A/B tức thì trong tích tắc
3 dB	Nhận thấy sự thay đổi rõ ràng khi so sánh	Gấp đôi công suất âm thanh chỉ mang lại mức tăng cảm nhận vừa phải này
5 dB	Thay đổi rất rõ rệt trong điều kiện nghe bình thường	Là bậc thang cảm nhận thực tế; tương ứng 6 bậc sắc thái âm nhạc cổ điển
10 dB	Cảm nhận âm thanh 'TO GẤP ĐÔI' hoặc 'GIẢM MỘT NỬA'	Mốc tâm lý âm học quan trọng khi cần tạo thay đổi âm lượng rõ rệt
Từ 15 dB trở lên	Đủ để tách biệt hoàn toàn hai lớp âm thanh khỏi nhau	Mức chênh lệch để lời thoại nổi hẳn trên nhạc nền

Mối liên hệ với 6 bậc sắc thái của âm nhạc cổ điển: một dàn nhạc giao hưởng biểu diễn trong dải động thực tế từ 60 dB SPL đến 90 dB SPL (biên độ biến thiên 30 dB). Chia cho bước nhảy cảm nhận thực tế 5 dB sẽ thu được đúng 6 bậc thang cường độ: Pianissimo (pp, rất êm, 60–65 dB) → Piano (p, êm nhẹ, 65–70 dB) → Mezzo-piano (mp, vừa êm, 70–75 dB) → Mezzo-forte (mf, vừa mạnh, 75–80 dB) → Forte (f, mạnh, 80–85 dB) → Fortissimo (ff, rất mạnh, 85–90 dB). Như

vậy, hệ thống ký âm mà các nhà soạn nhạc cổ điển đúc kết qua hàng trăm năm trùng khớp với quy luật tâm lý âm học hiện đại.

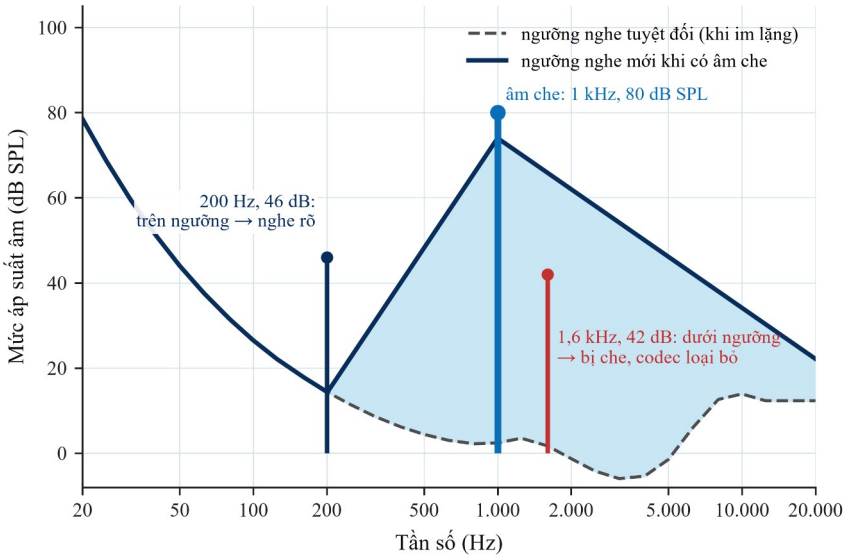
1.6.4 Hiện tượng che lấp âm thanh

Hiện tượng che lấp âm thanh (Auditory Masking) là hiện tượng một âm thanh có năng lượng mạnh hơn làm nâng ngưỡng nghe của thính giác, khiến một âm thanh yếu hơn xuất hiện cùng lúc hoặc kế cận trở nên hoàn toàn không nghe thấy được [17]:

1. **Che lấp đồng thời (Simultaneous Masking):** xảy ra khi hai âm thanh cùng tồn tại trong không gian. Đường cong che lấp có tính bất đối xứng rõ rệt: âm tần số thấp (âm trầm) che lấp âm tần số cao (âm bổng) dễ dàng và trên dải tần rộng hơn nhiều so với chiều ngược lại. Vì vậy tiếng trống kick hoặc tiếng rền động cơ xe tải dễ 'nuốt' giọng nói và tiếng đàn guitar nếu dải trầm không được xử lý.
2. **Che lấp theo thời gian (Temporal Masking):** gồm che lấp sau (Post-masking), kéo dài từ 50 đến 200 ms sau khi âm mạnh kết thúc, và che lấp trước (Pre-masking), kéo dài từ 5 đến 20 ms trước khi âm mạnh thực sự xuất hiện trong thính giác do não xử lý âm lớn nhanh hơn.

Hiện tượng che lấp có hai ứng dụng quan trọng trong công nghệ truyền thông hiện đại:

1. **Thuật toán nén âm thanh có tổn hao (Lossy Audio Codecs):** các thuật toán như MP3, AAC, Ogg Vorbis, Opus xây dựng một mô hình tâm lý âm học (Psychoacoustic Model) để tính chính xác thành phần phổ nào đang bị các âm lớn che lấp, rồi loại bỏ các bit dữ liệu tương ứng mà tai người không nhận ra, nhờ đó giảm kích thước tệp tới 10 lần.
2. **Kỹ thuật Sidechain Ducking:** tự động hạ âm lượng nền nhạc mỗi khi có lời thoại dựa trên nguyên lý che lấp; kỹ thuật này thường dùng trong phim tài liệu, phóng sự và podcast.

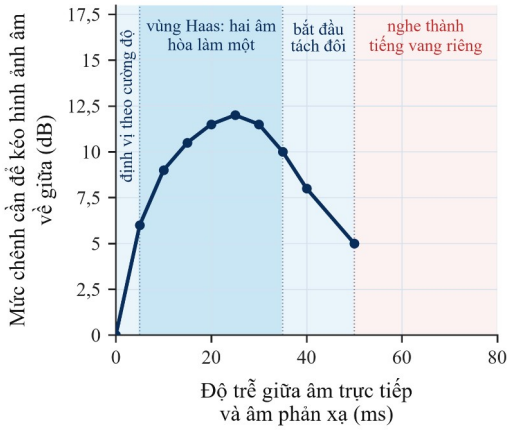


Hình 1.23 Đường cong ngưỡng che lấp trong miền tần số — cơ sở vật lý của mọi thuật toán nén có tổn hao

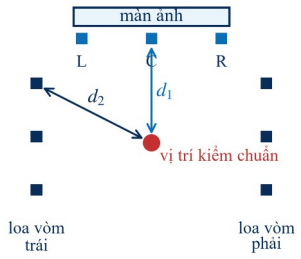
1.6.5 Cảm nhận không gian hai tai và hiệu ứng Haas

Não bộ con người định vị phương vị nguồn âm trong không gian 3 chiều dựa trên sự sai khác tín hiệu giữa hai tai (Binaural Hearing), gồm chênh lệch thời gian giữa hai tai (ITD, Interaural Time Difference, tối đa 0,6–0,7 ms, chủ đạo dưới 1,5 kHz) và chênh lệch mức áp suất giữa hai tai (ILD, Interaural Level Difference, do đầu che chắn tạo 'bóng âm', chủ đạo trên 1,5 kHz). Ngoài ra, các nếp gấp vành tai tạo nên hàm truyền liên quan đầu người (HRTF, Head-Related Transfer Function), giúp phân biệt trước/sau và độ cao.

Hiệu ứng Haas (hiệu ứng đến trước, Precedence Effect): nhà khoa học Helmut Haas đã chứng minh một quy luật tâm sinh lý thính giác kinh điển, theo đó khi hai hay nhiều nguồn phát cùng một tín hiệu âm thanh từ các hướng khác nhau và sóng âm đến tai người nghe với độ chênh lệch thời gian dưới 50 ms (0,05 giây), não bộ hợp nhất các sóng âm đó thành một sự kiện duy nhất và định vị nguồn âm theo hướng của âm thanh đến tai sớm nhất. Âm thanh đến sau chỉ có tác dụng làm tăng độ to và mở rộng cảm giác không gian [13].



(a) đường cong hiệu ứng ưu tiên



(b) cài độ trễ loa vòm rạp chiếu

Hình 1.24 Đường cong hiệu ứng ưu tiên (Haas) và sơ đồ cài đặt độ trễ cho hệ loa vòm

1.6.6 Tiêu chuẩn an toàn và bảo vệ thính giác nghề nghiệp theo NIOSH

Viện Quốc gia về An toàn và Sức khỏe Nghề nghiệp Hoa Kỳ (NIOSH) [7] ban hành bảng giới hạn thời gian tiếp xúc tiếng ồn an toàn tối đa dựa trên Quy tắc trao đổi 3 dB (3-dB Exchange Rule: mỗi khi mức áp suất âm tăng thêm 3 dBA thì thời gian tiếp xúc cho phép giảm đi một nửa):

Bảng 1.12 Tiêu chuẩn giới hạn thời gian tiếp xúc tiếng ồn an toàn nghề nghiệp (NIOSH)

Mức áp suất âm (dBA)	Thời gian tiếp xúc tối đa cho phép trong ngày (NIOSH)
85 dBA	8 giờ làm việc liên tục
88 dBA	4 giờ
91 dBA	2 giờ
94 dBA	1 giờ
97 dBA	30 phút

Mức áp suất âm (dBA)	Thời gian tiếp xúc tối đa cho phép trong ngày (NIOSH)
100 dBA	15 phút
103 dBA	7,5 phút
106 dBA	3,75 phút
112 dBA	Dưới 1 phút
140 dBA	0 giây (Tức thì)

1. **Nút tai lọc phẳng chuyên dụng:** người làm âm thanh nên luôn mang nút tai lọc phẳng dành cho nhạc sĩ và kỹ sư âm thanh (Musician's Earplugs). Khác với nút tai bọt xốp chi phí thấp, vốn cắt dải cao và làm méo âm thanh, nút tai lọc phẳng giảm đều 15–25 dB trên toàn dải tần, nhờ đó người làm âm thanh thẩm định chính xác bản mix mà vẫn bảo vệ được thính giác.
2. **Quy tắc 60/60 khi dùng tai nghe:** cần áp dụng nghiêm ngặt Quy tắc 60/60 khi làm việc với tai nghe kiểm âm (Headphones), tức là không nghe quá 60% mức âm lượng tối đa của thiết bị và không nghe liên tục quá 60 phút mà không nghỉ.
3. **Quản lý thời gian làm việc:** nên nghỉ ngơi đôi tai trong không gian yên tĩnh 10–15 phút sau mỗi 90 phút hòa âm liên tục để tránh hiện tượng dịch chuyển ngưỡng nghe tạm thời (TTS, Temporary Threshold Shift), nguyên nhân khiến kỹ sư đẩy thừa dải cao ở cuối buổi làm việc.

1.6.7 Âm thanh nhị nhĩ và công nghệ âm thanh theo đối tượng

Âm thanh nhị nhĩ (Binaural Audio) là kỹ thuật ghi âm hoặc tái tạo âm thanh mô phỏng chính xác những gì hai màng nhĩ con người nhận được trong không gian thực tế ba chiều (bao gồm ITD, ILD và HRTF). Tại hiện trường, các chuyên gia sử dụng thiết bị 'Đầu người giả' (Dummy Head, chẳng hạn Neumann KU 100) có vành tai cao su đúc nguyên bản và micro đặt đúng vị trí màng nhĩ để thu tiếng môi trường. Khi phát lại qua tai nghe, người nghe cảm nhận một không gian 3D sống động hơn so với định dạng Stereo thông thường.

Xu hướng công nghệ hiện đại: lĩnh vực âm thanh đang chuyển dịch từ tư duy kênh truyền thống (Channel-based: Stereo, 5.1, 7.1) sang tư duy âm thanh dựa trên đối tượng (Object-based Audio, điển hình là Dolby Atmos). Mỗi nguồn âm thanh là một 'đối tượng âm thanh độc lập' mang theo siêu dữ liệu tọa độ không gian 3 chiều (metadata X, Y, Z). Bộ giải mã âm thanh tại thiết bị phát tự động kết xuất (render) tín hiệu theo thời gian thực tới số lượng loa hoặc tai nghe thực tế có sẵn, đồng thời có thể kết hợp công nghệ theo dõi chuyển động đầu (Head Tracking) để giữ không gian âm thanh cố định trong không gian thực.

1.7 Mô phỏng kỹ thuật và cầu nối lập trình âm thanh số

Trong các hệ thống âm thanh số hiện đại, mọi tín hiệu âm thanh đều được số hóa thành các chuỗi mẫu số rời rạc (discrete digital samples) và xử lý thông qua các thuật toán xử lý tín hiệu số (DSP). Để người đọc kết nối trực quan các công

thức vật lý lý thuyết với cách máy tính và phần mềm DAW vận hành thực tế, dưới đây là chương trình mô phỏng viết bằng ngôn ngữ lập trình Python:

Liệt kê 1.3 ch1_audio_physics_sim.py — Mô phỏng vật lý âm học và đo lường hiện trường

```
# SÁCH KỸ THUẬT AUDIO - VIDEO
# CHƯƠNG 1: MÔ PHỎNG VẬT LÝ ÂM HỌC VÀ ĐO LƯỜNG ĐỘNG HỌC HIỆN TRƯỜNG
# Tệp tin thực thi: ch1_audio_physics_sim.py
# Môi trường yêu cầu: Python 3.8+ kết hợp thư viện numpy và scipy
#
================================================================================

import numpy as np

def generate_and_analyze_tone(freq_hz=440.0, duration_sec=1.0,
sample_rate=48000, amplitude=1.0):
    """
    Sinh tín hiệu sóng sin thuần khiết (Nốt La chuẩn A4 = 440 Hz)
    và đo lường các thông số biên độ: Peak, RMS, Crest Factor, quy đổi dBFS.
    """
    # 1. Tạo trục thời gian rời rạc theo tần số lấy mẫu fs = 48.000 Hz
    num_samples = int(sample_rate * duration_sec)
    t = np.linspace(0, duration_sec, num_samples, endpoint=False)

    # 2. Phương trình sóng áp suất sin điều hòa: p(t) = P_peak * sin(2*pi*f*t)
    signal = amplitude * np.sin(2.0 * np.pi * freq_hz * t)

    # 3. Đo lường các giá trị biên độ thực tế
    peak_val = np.max(np.abs(signal))
    rms_val = np.sqrt(np.mean(signal**2))

    # 4. Quy đổi sang thang đo Decibel số toàn dải (dBFS) với mốc trần 1.0 = 0
dBFS
    dbfs_peak = 20.0 * np.log10(peak_val)
    dbfs_rms = 20.0 * np.log10(rms_val)
    crest_factor_db = dbfs_peak - dbfs_rms

    print("=" * 70)
    print(f"Báo cáo phân tích tín hiệu sóng sin (f = {freq_hz} Hz, fs =
{sample_rate} Hz)")
    print("=" * 70)
    print(f"• Biên độ Đỉnh (Peak Amplitude)   : {peak_val:.6f}  --
>  {dbfs_peak:6.2f} dBFS")
    print(f"• Giá trị Hiệu dụng (RMS Level)   : {rms_val:.6f}  --
>  {dbfs_rms:6.2f} dBFS")
    print(f"• Hệ số Đỉnh (Crest Factor)      : {crest_factor_db:6.2f} dB
(Chuẩn lý thuyết: 3.01 dB)")

    # Kiểm tra nguy cơ méo cắt đỉnh (Clipping Check)
    if peak_val >= 1.0:
        print("  [Cảnh báo]: Tín hiệu chạm mức trần 0 dBFS! Nguy cơ méo cắt
đỉnh (Clipping).")
    else:
        headroom_db = -dbfs_peak
        print(f"  [An toàn]: Khoảng dự trữ an toàn (Headroom) khả dụng:
{headroom_db:.2f} dB")

    return signal

def calculate_inverse_square_law(spl_initial=80.0, d_initial=0.5,
d_target=1.0):
    """
```

```
    Tính toán mức suy hao áp suất âm thanh theo định luật bình phương nghịch
đảo
    áp dụng cho bài toán khoảng cách đặt micro.
    """
    delta_db = 20.0 * np.log10(d_initial / d_target)
    spl_final = spl_initial + delta_db

    print("-" * 70)
    print("Mô phỏng định luật bình phương nghịch đảo (khoảng cách micro)")
    print("-" * 70)
    print(f"• Cự ly ban đầu d1 = {d_initial:.2f} m : Mức thu nhận =
{spl_initial:.2f} dB SPL")
    print(f"• Cự ly dời xa d2 = {d_target:.2f} m : Mức thu nhận =
{spl_final:.2f} dB SPL")
    print(f"• Độ sụt giảm mức tín hiệu (ΔL)   : {delta_db:.2f} dB (Đúng -6.02
dB khi nhân đôi cự ly)")
    print(f"• Bù Gain tiền khuếch đại Preamp  : Bắt buộc tăng +{-delta_db:.2f}
dB (SNR bị giảm tương ứng!)")
    print("-" * 70)
    return spl_final


if __name__ == "__main__":
    # Thực thi phân tích sóng sin chuẩn nốt La 440 Hz với biên độ toàn phần
    audio_signal = generate_and_analyze_tone(freq_hz=440.0, duration_sec=1.0,
amplitude=1.0)

    # Mô phỏng bài toán dời micro từ 0.5m lên 1.0m
    spl_result = calculate_inverse_square_law(spl_initial=80.0, d_initial=0.5,
d_target=1.0)
```

Phân tích kết quả thuật toán mô phỏng: khi chạy đoạn mã trên, chương trình tính được giá trị RMS của sóng sin đạt 0,707107 và mức RMS tương ứng là -3,01 dBFS, nghĩa là độ chênh lệch giữa Peak và RMS bằng 3,01 dB, khớp với công thức lý thuyết. Đồng thời, hàm tính suy hao khoảng cách cho thấy mỗi khi cự ly tăng gấp đôi (từ 0,5 m lên 1,0 m), mức thanh áp giảm -6,02 dB, điều cần lưu ý khi chọn khoảng cách đặt micro.

TÓM TẮT CHƯƠNG 1

Nội dung cốt lõi của Chương 1 được đúc kết thành 15 luận điểm sau:

1. Bản chất sóng dọc: âm thanh là sóng cơ học lan truyền năng lượng dao động qua môi trường đàn hồi, và là sóng dọc trong chất khí. Sóng âm truyền năng lượng mà không truyền vật chất. Sự tồn tại của âm thanh đòi hỏi ba điều kiện: nguồn âm dao động, môi trường truyền dẫn và bộ phận tiếp nhận dao động.
2. Vận tốc và bước sóng: vận tốc âm trong không khí xấp xỉ $c \approx 331{,}3 + 0{,}606 \cdot \theta$ (m/s), không phụ thuộc vào áp suất và tần số ở điều kiện thường. Bước sóng $\lambda = c/f$ là cầu nối giữa con số và kích thước vật lý, quy định kích thước màng loa Subwoofer (vài mét) và khả năng che chắn, hấp thụ của vật liệu tiêu âm.
3. Trở kháng âm học đặc trưng $Z = \rho c$ quyết định tỷ lệ sóng âm truyền qua hay bị phản xạ tại mặt phân cách, và sự chênh lệch trở kháng giữa không khí và nước giải thích chức năng phối hợp trở kháng của chuỗi xương tai giữa.
4. Bốn đại lượng sóng sin cơ bản: biên độ (Peak, RMS, Crest Factor) quyết định âm lượng và ranh giới méo cắt clipping; tần số f (Hz) quyết định cao độ trên

thang logarit 10 quãng tám; âm sắc được tạo nên bởi chuỗi hài Fourier và đường bao ADSR (thông tin nhận diện nằm ở 50 ms đầu của giai đoạn Attack); góc pha quy định sự cộng dồn (+6 dB) hay triệt tiêu hoàn toàn.

5. Lọc lược (Comb Filtering) hình thành khi trộn hai tín hiệu lệch thời gian Δt, với các hõm triệt tiêu tại $f = (2n+1)/(2\Delta t)$. Quy tắc khoảng cách 3:1 ($d_{mic\text{-}mic} \geq 3 \cdot d_{nguồn\text{-}mic}$) là nguyên tắc quan trọng khi thu nhiều nguồn âm bằng nhiều micro.
6. Âm học phòng: cần phân biệt sóng trực tiếp, tiếng vang Echo (trễ ≥ 100 ms, cự ly ≥ 17 m) và âm vang Reverberation (trễ < 50 ms). Sóng đứng (Standing Waves) dải trầm tại các góc phòng dựng phải được xử lý bằng bẫy âm trầm (Bass Trap) và vị trí ngồi nghe 38%.
7. Thời gian âm vang RT60 tính theo công thức Sabine và Eyring. Tần số Schroeder $f_s \approx 2000\sqrt{(RT60/V)}$ phân định ranh giới giữa chế độ xử lý sóng đứng rời rạc (bằng bass trap) và chế độ âm học hình học thống kê (bằng tiêu âm, tán âm).
8. Bản chất đơn vị Decibel: decibel luôn là một tỷ số logarit, với hệ số 20 cho đại lượng biên độ trường (áp suất, điện áp) và hệ số 10 cho đại lượng công suất, năng lượng. Hai nguồn không tương quan cùng mức cộng lại tăng +3 dB, còn hai tín hiệu tương quan cùng pha tăng +6 dB.
9. Định luật bình phương nghịch đảo: trong trường tự do, mức áp suất âm thanh suy giảm 6,02 dB mỗi khi nhân đôi khoảng cách. Bài học rút ra là không bù khoảng cách đặt micro bằng Gain, vì làm như vậy sẽ giảm nghiêm trọng tỷ số tín/tạp SNR.
10. Đo lường độ to hiện đại theo ITU-R BS.1770: đơn vị LUFS/LKFS tích hợp mạch lọc trọng số K và cơ chế đóng cổng Gating. Mức trần True Peak (dBTP) bảo vệ tín hiệu khỏi méo liên mẫu khi mã hóa phát hành trực tuyến (EBU R128: -23 LUFS, YouTube: -14 LUFS, Netflix: -27 LKFS).
11. Gain Staging chuẩn mực trong DAW: duy trì mức làm việc danh định của lời thoại ở -18 dBFS RMS đến -20 dBFS RMS để chừa khoảng dự trữ Headroom an toàn chống clipping.
12. Cơ chế phát âm giọng nói hoạt động theo mô hình Nguồn – Bộ lọc (Source-Filter Model): hộp cộng hưởng khoang miệng và mũi tạo nên các dải Formants (F1, F2, F3) nhận diện nguyên âm và chất giọng. Kỹ thuật khoét tần số Frequency Carving EQ giúp tách lời thoại khỏi nhạc nền.
13. Sinh lý thính giác: ống tai ngoài cộng hưởng khuếch đại tự nhiên dải 2,5–4 kHz, còn tai trong chứa ốc tai với màng đáy phân tích phổ cơ học tonotopic. Tế bào lông mao không tái sinh, nên thính giác phải được bảo vệ theo tiêu chuẩn NIOSH (quy tắc trao đổi 3 dB, nút tai lọc phẳng).
14. Đường cong đẳng âm ISO 226 (Fletcher-Munson): đáp tuyến tai thay đổi theo mức âm lượng, vì vậy hệ thống loa kiểm âm cần được hiệu chuẩn ở mức chuẩn 85 dBC SPL (hoặc 79 dBC SPL cho phòng dựng nhỏ). Ngưỡng chênh lệch: 1 dB (tối thiểu), 5 dB (bậc cảm nhận rõ), 10 dB (to gấp đôi), 15 dB (lời thoại nổi trên nền).

15. Hiện tượng che lấp âm thanh là cơ sở của các codec nén MP3/AAC/Opus và sidechain ducking. Hiệu ứng Haas (< 50 ms) là cơ sở cài đặt độ trễ (Delay) cho hệ thống loa vòm để giữ định vị âm thanh theo hướng loa chính.

1.8 Hệ thống câu hỏi và bài tập

1.8.1 Câu hỏi ôn tập lý thuyết

1. Nêu rõ ba điều kiện tiên quyết cho sự hình thành và tiếp nhận âm thanh. Vì sao trong thực tế ngoài vũ trụ không thể có tiếng nổ truyền qua chân không?
2. Phân biệt sự khác nhau cơ bản giữa sóng dọc và sóng ngang. Trong không khí, sóng âm thuộc loại sóng nào và điều đó có ý nghĩa gì đối với việc thiết kế bộ chống rung (Shockmount) cho giá micro?
3. Chứng minh vì sao sóng âm truyền năng lượng chứ không truyền vật chất. Nêu độ lớn biên độ dịch chuyển của phân tử không khí ở mức 94 dB SPL và tại ngưỡng nghe 0 dB SPL.
4. Vì sao vận tốc âm trong không khí không phụ thuộc áp suất khí quyển nhưng lại thay đổi theo nhiệt độ môi trường? Nêu hệ quả thực tế khi tính toán độ trễ cho hệ thống âm thanh đại nhạc hội ngoài trời giữa buổi trưa và đêm.
5. Trở kháng âm học đặc trưng $Z = \rho c$ là gì? Sử dụng khái niệm này để giải thích vì sao một người lặn dưới nước gần như không nghe được tiếng nói của người đứng trên bờ.
6. Phân biệt ba khái niệm đo lường biên độ: Biên độ đỉnh (Peak), Đỉnh–đỉnh (Peak-to-Peak) và Giá trị hiệu dụng (RMS). Hệ số đỉnh (Crest Factor) của một bản mix nói lên điều gì về mức độ nén ép độ động?
7. Định nghĩa bước sóng $\lambda = c/f$. Dựa trên khái niệm bước sóng, giải thích vì sao loa siêu trầm Subwoofer cần kích thước màng loa lớn (15–18 inch) trong khi loa tép Tweeter lại nhỏ.
8. Phân biệt sự khác nhau giữa Đảo cực (Polarity Inversion) và Lệch pha (Phase Shift). Nút 'Ø' trên bàn mixer xử lý được vấn đề nào và không xử lý được vấn đề nào?
9. Trình bày bản chất vật lý của hiện tượng lọc lược (Comb Filtering). Nêu công thức xác định các tần số bị triệt tiêu và giải thích vì sao một bản mix Stereo nghe tốt lại có thể bị 'sập tiếng' khi chuyển sang Mono.
10. Trình bày nội dung và ý nghĩa của Quy tắc khoảng cách 3:1 khi bố trí micro thu âm hai người đối thoại.
11. Phân biệt sự khác nhau về bản chất vật lý và cảm giác thính giác giữa Sóng trực tiếp, Tiếng vang (Echo) và Âm vang (Reverberation). Điều kiện cự ly tối thiểu để hình thành tiếng vang trong phòng kín là bao nhiêu mét?
12. Sóng đứng (Standing Waves) trong phòng kín hình thành như thế nào? Phân tích nguy cơ kỹ thuật nghiêm trọng đối với người dựng phim nếu vô tình đặt bàn dựng đúng vào Nút sóng tại chính giữa phòng.
13. Trình bày công thức Sabine tính thời gian âm vang RT60. Khi nào bắt buộc phải chuyển sang sử dụng công thức Eyring?

14. Ý nghĩa vật lý của Tần số Schroeder (f_s) là gì? Nêu chiến lược xử lý âm học cho vùng tần số dưới f_s và trên f_s.
15. Phân tích hiện tượng khúc xạ âm thanh do nghịch nhiệt khí quyển vào ban đêm. Điều này ảnh hưởng thế nào đến việc thu âm hoặc phát loa ngoài trời vào ban đêm gần khu dân cư?
16. Trình bày Hiệu ứng Doppler và nêu một ứng dụng trong thiết kế âm thanh (Sound Design) cho nguồn âm chuyển động.
17. Giải thích vì sao công thức decibel dùng hệ số 20 với đại lượng biên độ trường (áp suất, điện áp) nhưng dùng hệ số 10 với đại lượng công suất, năng lượng, đồng thời chứng minh mối liên hệ giữa hai hệ số này.
18. Hai máy phát điện giống hệt nhau, mỗi máy gây mức ồn 80 dB SPL. Hãy xác định mức ồn tổng khi cả hai máy cùng hoạt động và giải thích vì sao kết quả không phải là 160 dB.
19. Trình bày bốn giai đoạn xử lý cốt lõi của thuật toán đo độ to tiêu chuẩn ITU-R BS.1770, đồng thời cho biết cơ chế đóng cổng (gating) giải quyết vấn đề gì đối với các đoạn tĩnh lặng trong chương trình.
20. Phân biệt Sample Peak (dBFS) với True Peak (dBTP), đồng thời giải thích vì sao một tệp âm thanh có Sample Peak đạt 0 dBFS vẫn có thể gây méo tiếng nghiêm trọng khi mã hóa để phát hành trên YouTube.
21. Phát biểu định luật bình phương nghịch đảo và phân tích ý nghĩa của định luật này khi micro bị dời xa nguồn âm từ 0,5 m lên 1,0 m.
22. Nêu khái niệm formant, giải thích cơ chế tạo nên các dải formant của giọng nói con người và trình bày kỹ thuật khoét tần số (Frequency Carving EQ) để tách lời thoại khỏi nhạc nền.
23. Mô tả cấu tạo ba khoang tai người, giải thích cơ chế phối hợp trở kháng của chuỗi xương tai giữa và cho biết vì sao tai người nhạy cảm nhất ở dải tần 2,5–4 kHz.
24. Trình bày ý nghĩa của hệ thống đường cong đẳng âm ISO 226 và giải thích vì sao cần hiệu chuẩn hệ thống loa kiểm âm ở một mức chuẩn cố định (85 dBC SPL).
25. Trình bày hiệu ứng Haas (hiệu ứng đến trước) và giải thích vì sao phải cài đặt độ trễ kỹ thuật số (Digital Delay) cho hệ thống loa vòm đặt gần người nghe hơn loa chính.

1.8.2 Bài tập tình huống có hướng dẫn giải

Bài tập tình huống 1: Cự ly đặt micro và sự suy giảm tỷ số tín hiệu trên tạp âm (SNR)

Đề bài: Trong một buổi thu âm ngoài trời ở nhiệt độ 20°C, micro được đặt cách miệng người nói 0,6 m và máy đo ghi nhận mức áp suất âm thanh lời thoại là 76 dB SPL. Ở lần thu tiếp theo, micro phải dời ra cự ly 2,4 m so với người nói.

a) Mức áp suất âm thanh lời thoại thu được tại micro ở cự ly 2,4 m là bao nhiêu dB SPL?

b) Để giữ nguyên mức điện áp tín hiệu cấp vào máy ghi âm, người thu âm phải tăng Gain tiền khuếch đại thêm bao nhiêu dB? Hãy phân tích tác hại của việc tăng Gain này đối với chất lượng bản thu thoại.

Hướng dẫn giải chi tiết: a) Tỷ số khoảng cách $r_2 / r_1 = 2,4 / 0,6 = 4$ lần. Khoảng cách tăng 4 lần tương đương 2 lần nhân đôi cự ly (từ 0,6 m → 1,2 m → 2,4 m).

Áp dụng định luật bình phương nghịch đảo: $\Delta L = 20 \cdot \log_{10}(r_1 / r_2) = 20 \cdot \log_{10}(0,6 / 2,4) = 20 \cdot \log_{10}(1/4) = -20 \cdot \log_{10}(4) = -20 \times 0,60206 \approx -12,04$ dB.

Mức áp suất âm thanh thu được tại cự ly 2,4 m: $L_2 = 76$ dB - 12,04 dB ≈ 63,96 dB SPL (làm tròn thành 64 dB SPL).

b) Để giữ nguyên mức điện áp đầu ra cấp cho bộ biến đổi A/D, kỹ sư phải tăng núm Gain của tiền khuếch đại (Preamp) thêm đúng +12 dB.

Phân tích tác hại nghề nghiệp: khi tăng Gain thêm 12 dB, toàn bộ tiếng gió, tiếng ồn phố xá xung quanh và đặc biệt là tiếng nhiễu nhiệt điện tử (EIN) của chính mạch micro cũng bị khuếch đại thêm đúng 12 dB. Do đó, tỷ số tín hiệu trên tạp âm (SNR) của bản thu thoại giảm nghiêm trọng tới 12 dB, khiến tiếng nói chìm trong nhiễu nền và rất khó làm sạch ở khâu xử lý sau mà không gây méo tiếng.

Bài tập tình huống 2: Tính toán sóng đứng và mode trục trong phòng dựng phim

Đề bài: Một phòng dựng phim có kích thước hình hộp chữ nhật với chiều dài L = 5,0 m, chiều rộng W = 3,5 m và chiều cao H = 2,8 m. Nhiệt độ không khí duy trì ở 20°C (vận tốc âm c = 343 m/s).

a) Hãy tính tần số sóng đứng đầu tiên (chế độ trục bậc 1, Axial Mode f_1) dọc theo chiều dài L của căn phòng.

b) Tần số này nằm trong phân vùng âm thanh nào? Nếu kỹ sư dựng phim kê bàn làm việc ở chính giữa phòng theo chiều dài (tọa độ x = 2,5 m), hiện tượng âm học nào sẽ xảy ra và gây tác hại thế nào đến bản hòa âm?

Hướng dẫn giải chi tiết: a) Điều kiện hình thành sóng đứng chế độ trục bậc 1 ($n_x = 1$, $n_y = 0$, $n_z = 0$) giữa hai bức tường cách nhau L là $L = \lambda / 2$, suy ra bước sóng $\lambda = 2 \cdot L = 2 \times 5{,}0 = 10{,}0$ m.

Tần số cộng hưởng trục bậc 1: $f_1 = c / \lambda = c / (2 \cdot L) = 343 / (2 \times 5{,}0) = 34{,}3$ Hz.

b) Phân tích chuyên môn: Tần số 34,3 Hz nằm trong phân vùng siêu trầm (Sub-bass: 20–60 Hz). Sóng đứng ở tần số này tạo ra bụng sóng tích tụ năng lượng áp suất cực đại tại hai bức tường phía trước và phía sau, đồng thời tạo ra nút sóng triệt tiêu áp suất hoàn toàn tại chính giữa phòng (x = L/2 = 2,5 m).

Nếu kỹ sư ngồi làm việc đúng chính giữa phòng, đôi tai của họ nằm trọn trong nút sóng của tần số 34,3 Hz và các hài âm bậc lẻ kế tiếp nên tiếng bass sâu biến mất hoàn toàn. Kỹ sư dễ nhầm rằng bản thu thoại hoặc tiếng động bị thiếu trầm và tự ý đẩy EQ bass lên thêm +6 đến +10 dB. Khi bản mix được phát trên hệ thống loa có subwoofer, tiếng bass sẽ bị thừa quá mức, gây rung và ù, làm hỏng chất lượng bản hòa âm. Giải pháp là dời vị trí ngồi nghe về tọa độ 38% chiều dài phòng (x ≈ 1,9 m) và lắp đặt các khối bẫy âm trầm (Bass Trap) tại 4 góc phòng.

Bài tập tình huống 3: Cài đặt độ trễ loa vòm theo hiệu ứng Haas

Đề bài: Trong một phòng nghe lớn, khoảng cách từ cụm loa chính phía trước (Center) đến vị trí nghe trung tâm là dmàn = 18,0 m, trong khi dàn loa vòm gắn trên vách tường bên (Surround Speakers) chỉ cách vị trí này dvòm = 3,0 m. Vận tốc truyền âm trong phòng là c = 340 m/s.

a) Tính thời gian sóng âm truyền từ loa chính và từ loa vòm tới tai người nghe nếu hai loa phát tín hiệu đồng thời.

b) Để kích hoạt hiệu ứng Haas, giúp người nghe định vị âm thanh theo hướng loa chính, đồng thời âm thanh loa vòm đến sau khoảng 20 ms, cần cài đặt thời gian trễ kỹ thuật số (Digital Delay) trên bộ DSP cấp cho kênh loa vòm là bao nhiêu mili-giây (ms)?

Hướng dẫn giải chi tiết: a) Thời gian truyền âm tự nhiên từ từng cụm loa tới tai người nghe như sau.

• **Từ loa chính:** $t_{màn} = d_{màn} / c = 18,0 / 340 \approx 0,05294$ giây = 52,94 ms.

• **Từ loa vòm gắn tường:** $t_{vòm} = d_{vòm} / c = 3,0 / 340 \approx 0,00882$ giây = 8,82 ms.

Nhận xét: nếu không cài đặt độ trễ, sóng âm từ loa vòm gắn tường sẽ đến tai người nghe sớm hơn sóng âm từ loa chính: 52,94 - 8,82 = 44,12 ms, vì vậy người nghe sẽ cảm nhận âm thanh phát ra từ bức tường bên cạnh.

b) Áp dụng nguyên lý hiệu ứng Haas: để não bộ định vị nguồn âm xuất phát từ loa chính, âm thanh từ loa chính phải tới tai trước tiên (tại mốc 52,94 ms) và âm thanh từ loa vòm phải tới sau đó 20 ms (nằm trong vùng Haas 5–30 ms).

Thời điểm mong muốn sóng âm loa vòm tới tai người nghe: $t_mong_muốn = t_{màn} + 20$ ms = 52,94 + 20,0 = 72,94 ms.

Thời gian trễ điện tử Digital Delay cần cài đặt trên bộ xử lý tín hiệu DSP được tính như sau.

$Delay_DSP = t_mong_muốn - t_{vòm}$ = 72,94 - 8,82 = 64,12 ms (cài đặt thực tế: 64 ms).

Đáp số: độ trễ cần cài đặt cho kênh loa vòm là 64,1 ms.

1.8.3 Bài tập tính toán định lượng kỹ thuật

Bài 1. Tính vận tốc truyền âm trong không khí ngoài trời ở nhiệt độ 28°C. → Hướng dẫn & Đáp số: $c \approx 331,3 + 0,606 \times 28 = 348,27$ m/s ≈ 348,3 m/s.

Bài 2. Ở nhiệt độ phòng thu tiêu chuẩn 20°C (c = 343 m/s), hãy tính bước sóng của các tần số âm thanh 250 Hz và 8.000 Hz (8 kHz), sau đó nhận xét về độ dày vật liệu tiêu âm cần thiết cho từng tần số. → Hướng dẫn & Đáp số: $\lambda_{250} = 343 / 250 = 1,372$ m; $\lambda_{8k} = 343 / 8.000 = 0,04288$ m ≈ 4,29 cm. Tiêu âm 250 Hz cần vật liệu dày tối thiểu λ/4 ≈ 34 cm hoặc bẫy màng, trong khi 8 kHz chỉ cần lớp mút mỏng khoảng 1–2 cm.

Bài 3. Một microphone thu tiếng động đặt cách nguồn phát 0,5 m được dời ra cự ly 2,0 m trong trường tự do. Hãy tính mức áp suất âm thanh thu được thay đổi bao nhiêu dB. → Hướng dẫn & Đáp số: Tỷ số $r_2/r_1 = 2,0 / 0,5 = 4$ lần. $\Delta L = 20 \cdot \log_{10}(0,5 / 2,0) = 20 \cdot \log_{10}(1/4) = -12,04$ dB (mức âm suy giảm đúng 12,04 dB).

Bài 4. Áp suất âm thanh đo được tại hiện trường là p = 0,2 Pa. Hãy tính mức áp suất âm thanh theo đơn vị dB SPL (biết $p_0 = 2 \cdot 10^{-5}$ Pa). → Hướng dẫn & Đáp số: $L_p = 20 \cdot \log_{10}(0,2 / 2 \cdot 10^{-5}) = 20 \cdot \log_{10}(10.000) = 20 \times 4 = 80$ dB SPL.

Bài 5. Ba nguồn ồn độc lập không tương quan, mỗi nguồn gây ra mức ồn 70 dB SPL, cùng hoạt động đồng thời. Hãy tính mức áp suất âm thanh tổng thể. → Hướng dẫn & Đáp số: $L_{tổng} = 10 \cdot \log_{10}(3 \times 10^{(70/10)}) = 10 \cdot \log_{10}(3 \times 10^7) = 70 + 10 \cdot \log_{10}(3) = 70 + 4,77 = 74,77$ dB SPL ≈ 74,8 dB SPL.

Bài 6. Hai nguồn ồn không tương quan có mức áp suất âm lần lượt là 80 dB SPL và 74 dB SPL. Hãy tính mức âm tổng và nhận xét về mức đóng góp của nguồn ồn yếu hơn. → Hướng dẫn & Đáp số: $L_{tổng} = 10 \cdot \log_{10}(10^{(80/10)} + 10^{(74/10)}) = 10 \cdot \log_{10}(100.000.000 + 25.118.864) = 10 \cdot \log_{10}(125.118.864) \approx 80,97$ dB SPL ≈ 81,0 dB SPL. Nguồn ồn 74 dB (nhỏ hơn 6 dB) chỉ làm tăng mức tổng thêm đúng 1,0 dB.

Bài 7. Hai bản sao của một tín hiệu thu thoại được trộn với nhau với độ trễ thời gian Δt = 0,5 ms (0,0005 s), gây hiện tượng lọc lược. Hãy tính ba tần số bị triệt tiêu hoàn toàn đầu tiên (f_{null}). → Hướng dẫn & Đáp số: Áp dụng công thức $f_{null} = (2n + 1) / (2 \cdot \Delta t)$. Với n = 0: $f_0 = 1 / (2 \times 0,0005) = 1.000$ Hz (1 kHz); với n = 1: $f_1 = 3 / 0,001 = 3.000$ Hz (3 kHz); với n = 2: $f_2 = 5 / 0,001 = 5.000$ Hz (5 kHz).

Bài 8. Một phòng dựng phim có chiều dài L = 6,2 m (c = 343 m/s). Hãy tính ba tần số sóng đứng chế độ trục đầu tiên theo chiều dài căn phòng. → Hướng dẫn & Đáp số: Áp dụng $f_n = n \cdot c / (2 \cdot L)$. Mode bậc 1 (n=1): $f_1 = 343 / (2 \times 6,2) = 27,66$ Hz ≈ 27,7 Hz; Mode bậc 2 (n=2): $f_2 = 2 \times 27,66 = 55,32$ Hz ≈ 55,3 Hz; Mode bậc 3 (n=3): $f_3 = 3 \times 27,66 = 82,98$ Hz ≈ 83,0 Hz.

Bài 9. Một phòng thu thanh có thể tích V = 120 m³ và tổng diện tích hấp thụ âm tương đương của các bề mặt là A = 40 sabin. Hãy tính thời gian âm vang RT60 theo công thức Sabine. → Hướng dẫn & Đáp số: $RT60 = 0,161 \cdot V / A = 0,161 \times 120 / 40 = 19,32 / 40 = 0,483$ giây ≈ 0,48 s.

Bài 10. Một cụm loa sân khấu có độ nhạy công bố 94 dB SPL @ 1 W / 1 m được cấp công suất điện P = 200 W. Hãy tính mức áp suất âm thanh tại vị trí khán giả cách loa 15 m trong điều kiện trường tự do ngoài trời. → Hướng dẫn & Đáp số: Mức âm tại 1 m khi cấp 200 W: $L_1m = 94 + 10 \cdot \log_{10}(200) = 94 + 23,01 = 117,01$ dB SPL. Mức suy hao tại cự ly 15 m: $\Delta L = 20 \cdot \log_{10}(15) = 20 \times 1,176 = 23,52$ dB. Mức SPL tại 15 m: $L_15m = 117,01 - 23,52 \approx 93,49$ dB SPL ≈ 93,5 dB SPL.

Bài 11. Một tập podcast hoàn thiện có mức Integrated Loudness là -21 LUFS và Max True Peak là -7,2 dBTP. Người sản xuất muốn đưa bản mix về chuẩn phát hành AES TD1004 (-16 LUFS, trần True Peak Max -1,0 dBTP). Hãy tính lượng Gain cần bù và kiểm tra xem mức True Peak sau khi nâng Gain có đạt ngưỡng an toàn hay không. → Hướng dẫn & Đáp số: Lượng Gain cần tăng thêm: ΔGain = (-16) - (-21) = +5,0 dB. Mức True Peak mới sau khi tăng Gain: True Peak_mới = -7,2 dBTP + 5,0 dB = -2,2 dBTP. Do -2,2 dBTP thấp hơn ngưỡng trần an toàn -1,0 dBTP, tệp đạt chuẩn phát hành mà không cần True-Peak Limiter can thiệp sâu.

Bài 12. Một đoạn trích tiếng động Foley có biên độ đỉnh đo được là A_{peak} = 0,90 và giá trị hiệu dụng A_{RMS} = 0,11 trên thang tuyến tính kỹ thuật số. Hãy tính hệ số đỉnh (Crest Factor) và nhận xét về mức độ nén dải động của đoạn tiếng động này.
→ Hướng dẫn & Đáp số: Crest Factor = 20 · $\log_{10}$(0,90 / 0,11) = 20 · $\log_{10}$(8,1818) = 20 × 0,91285 ≈ 18,26 dB ≈ 18,3 dB. Nhận xét: hệ số đỉnh 18,3 dB phản ánh một bản thu có dải động tự nhiên, âm thanh thoáng, chưa bị nén bởi compressor/limiter.

Bài 13. Cụm tháp loa tăng cường (Delay Tower) đặt cách sân khấu chính 60 m, nhiệt độ không khí ngoài trời đo được là 30°C. Hãy tính thời gian trễ Digital Delay cần cài đặt cho tháp loa này, biết rằng kỹ sư áp dụng thêm 12 ms độ trễ ưu tiên theo hiệu ứng Haas (Haas offset).
→ Hướng dẫn & Đáp số: Vận tốc âm ở 30°C: c = 331,3 + 0,606 × 30 = 349,48 m/s ≈ 349,5 m/s. Thời gian âm truyền tự nhiên: t = d / c = 60 / 349,48 ≈ 0,17168 s = 171,68 ms. Thời gian trễ cần cài đặt: Delay = 171,68 ms + 12 ms = 183,68 ms ≈ 184 ms.

Bài 14. Không khí ở 20°C có khối lượng riêng ρ = 1,204 kg/m³ và vận tốc truyền âm c = 343,4 m/s. Hãy tính trở kháng âm học đặc trưng Z của không khí và so sánh với trở kháng của nước ngọt ($\rho_{nước}$ = 1.000 kg/m³, $c_{nước}$ = 1.480 m/s).
→ Hướng dẫn & Đáp số: $Z_{khí}$ = ρ · c = 1,204 × 343,4 ≈ 413,45 rayl ≈ 413 rayl (Pa·s/m). $Z_{nước}$ = 1.000 × 1.480 = 1.480.000 rayl = 1,48 × 10^6 rayl. Tỷ lệ $Z_{nước}$ / $Z_{khí}$ ≈ 3.580 lần.

Bài 15. Một chiếc xe cứu thương phát còi hụ tần số f = 800 Hz đang chạy với vận tốc v_s = 20 m/s (72 km/h). Tính tần số âm thanh mà người quay phim đứng yên bên lề đường nghe được: khi xe đang tiến lại gần và khi xe đã chạy ra xa (cho biết vận tốc âm trong không khí là c = 343 m/s).
→ Hướng dẫn & Đáp số: Áp dụng công thức Doppler: Khi xe tiến lại gần: f' = f · c / (c - v_s) = 800 × 343 / (343 - 20) = 800 × 343 / 323 ≈ 849,54 Hz ≈ 849,5 Hz. Khi xe chạy ra xa: f'' = f · c / (c + v_s) = 800 × 343 / (343 + 20) = 800 × 343 / 363 ≈ 755,92 Hz ≈ 755,9 Hz.

1.8.4 Bài tự thực hành khảo sát âm học không gian thực tế

Mục tiêu: rèn luyện kỹ năng khảo sát âm học tại hiện trường và tự thẩm định kỹ thuật cho một không gian làm việc thực tế (phòng thu âm, phòng dựng hoặc góc làm việc cá nhân của người đọc).

1. **Nhiệm vụ 1 (đo đạc hình học và tính room modes):** sử dụng thước laser hoặc thước dây đo chính xác chiều dài (L), chiều rộng (W) và chiều cao (H) của phòng; tính thể tích V (m³) và diện tích các bề mặt bao quanh. Áp dụng công thức Rayleigh để tính ba tần số cộng hưởng chế độ trục đầu tiên theo từng chiều không gian và tính tần số Schroeder f_s.
2. **Nhiệm vụ 2 (đo kiểm tiếng ồn nền và đánh giá NC):** sử dụng máy đo mức âm cầm tay (hoặc ứng dụng đo SPL đã hiệu chuẩn trên điện thoại) để đo mức ồn nền dBA ở ba thời điểm khác nhau trong ngày (sáng, trưa, tối); sau đó lập

bảng thống kê và phân tích nguồn ồn chính (máy điều hòa, quạt máy tính, tiếng xe ngoài đường).

3. **Nhiệm vụ 3 (ước lượng RT60 hiện trạng):** lập bảng thống kê chất liệu thực tế của sàn, trần, tường và cửa kính; tra cứu Bảng 1.5 để tính tổng diện tích hấp thụ tương đương A (sabin) và tính RT60 theo công thức Sabine; đối chiếu với Bảng 1.6 để nhận xét phòng có đạt chuẩn hay không.
4. **Nhiệm vụ 4 (thu thử nghiệm và lập dự toán xử lý âm học):** thu một đoạn đọc kịch bản 30 giây bằng micro chuyên dụng và đo Integrated LUFS cùng Max True Peak trong DAW. Với ngân sách cải tạo giả định 5.000.000 đồng, lập phương án mua sắm vật liệu tiêu âm (bông khoáng Rockwool, rèm nhung, bẫy trầm) và thuyết minh phương án bố trí tối ưu.

1.9 Bảng đối chiếu thuật ngữ chuyên ngành Anh – Việt

Bảng 1.13 Thuật ngữ chuyên ngành âm thanh Anh – Việt

Thuật ngữ tiếng Anh (English Term)	Thuật ngữ tiếng Việt quy chuẩn	Định nghĩa chuyên môn & Ngữ cảnh sử dụng
Absorption coefficient	Hệ số hấp thụ	α từ 0 đến 1, phụ thuộc tần số
Acoustic Impedance	Trở kháng âm học	Z = ρc (đơn vị rayl); đại lượng quyết định mức độ phản xạ/truyền qua tại mặt phân cách
Acoustic Source	Nguồn phát âm	Vật thể dao động cơ học quanh vị trí cân bằng tạo sóng áp suất âm thanh
Acoustic wave	Sóng âm học	Sóng cơ học lan truyền qua môi trường vật chất đàn hồi
ADSR Envelope	Đường bao biên độ ADSR	Mô hình 4 giai đoạn biến thiên biên độ theo thời gian: Attack, Decay, Sustain, Release
Amplitude	Biên độ	Peak, peak-to-peak, RMS
Amplitude (Peak / RMS)	Biên độ (Đỉnh / Hiệu dụng)	Độ lệch áp suất cực đại tức thời và giá trị căn bậc hai trung bình bình phương năng lượng
Auditory Masking	Hiện tượng che lấp âm thanh	Âm thanh mạnh hơn nâng ngưỡng nghe làm che khuất các âm thanh yếu hơn lân cận
Axial Mode	Chế độ sóng đứng trục	Mode sóng đứng mạnh nhất hình thành giữa hai bề mặt phẳng song song đối diện nhau
Bass Trap	Bẫy âm trầm	Khối tiêu âm bông khoáng tỷ trọng cao đặt góc phòng triệt tiêu sóng đứng dưới 150 Hz
Bass trap / Diffuser	Bẫy trầm / Tấm tán âm	Hai nhóm giải pháp âm học khác nhau
Binaural Audio	Âm thanh nhị nhĩ	Kỹ thuật ghi/tổng hợp âm thanh 3D mô phỏng chính xác tín hiệu tại hai màng nhĩ tai người

Thuật ngữ tiếng Anh (English Term)	Thuật ngữ tiếng Việt quy chuẩn	Định nghĩa chuyên môn & Ngữ cảnh sử dụng
Clipping	Méo cắt đỉnh tín hiệu	Hiện tượng méo phi tuyến dạng sóng khi biên độ vượt quá mức trần số tuyệt đối 0 dBFS
Clipping distortion	Méo cắt đỉnh	Hiện tượng bẹt đầu sóng khi biên độ vượt ngưỡng trần số
Comb Filtering	Hiệu ứng lọc lược	Hiện tượng đáp tuyến tần số bị răng cưa (đỉnh +6 dB, hõm -∞ dB) do trộn tín hiệu lệch pha
Compression / Rarefaction	Vùng nén / vùng giãn	Hai pha của sóng âm
Crest Factor	Hệ số đỉnh tín hiệu	Tỷ số giữa biên độ Đỉnh và giá trị Hiệu dụng RMS tính bằng dB; phản ánh độ động bản phối
Critical band	Băng thông tới hạn	Thang Bark, thang ERB
Decibel (dB)	Đơn vị Đê-xi-ben	Tỷ số logarit thập phân đo lường mức độ âm thanh; dùng hệ số 20 (biên độ) và hệ số 10 (công suất)
Decibel relative to Full Scale (dBFS)	Decibel toàn dải số	Thang đo logarit lấy biên độ trần 1.0 làm mốc 0 dBFS
Decibel SPL (dB SPL)	Mức áp suất âm thanh	Thang đo áp suất âm thanh trong không khí quy chiếu theo ngưỡng nghe 20 µPa ($2 \cdot 10^{-5}$ Pa)
Digital Audio Workstation (DAW)	Trạm làm việc âm thanh số	Phần mềm chuyên dụng ghi âm, dựng, hòa âm và master (Fairlight, Pro Tools, Nuendo)
Direct Sound	Sóng âm trực tiếp	Sóng âm truyền thẳng theo đường ngắn nhất từ nguồn phát đến micro không qua phản xạ
Doppler Effect	Hiệu ứng Đốp-lơ	Sự dịch chuyển tần số cảm nhận khi nguồn âm và người nghe chuyển động tương đối so với nhau
Dynamic Range	Dải động	Khoảng cách tính bằng dB từ sàn nhiễu (Noise Floor) lên tới mức tín hiệu cực đại trước khi méo
Echo	Tiếng vang	Sóng âm phản xạ từ bề mặt xa trễ trên 100 ms tạo thành âm thanh rời rạc tách biệt
Envelope (ADSR)	Đường bao biên độ	Attack – Decay – Sustain – Release
Equal-loudness contour	Đường cong đẳng âm	ISO 226
Equal-Loudness Contours	Đường cong đẳng âm	Đồ thị biểu diễn mức SPL cần thiết theo tần số để tạo cùng độ to (ISO 226 / Fletcher-Munson)
Foley	Tiếng động điện ảnh hậu kỳ	Nghệ thuật thu tạo tiếng động mô phỏng chân thực hành động nhân vật trong phòng thu

Thuật ngữ tiếng Anh (English Term)	Thuật ngữ tiếng Việt quy chuẩn	Định nghĩa chuyên môn & Ngữ cảnh sử dụng
Formant	Formant, đỉnh cộng hưởng giọng	F1, F2 quyết định nguyên âm
Formants	Dải tần số tạo âm / Formant	Các đỉnh cộng hưởng khuếch đại của khoang miệng/mũi đặc trưng cho nguyên âm và chất giọng
Frequency (Hz) / Pitch	Tần số / Cao độ	Số chu kỳ dao động trong 1 giây (đại lượng vật lý) và cảm nhận trầm bổng (tâm lý âm học)
Frequency / Pitch	Tần số / Cao độ	Đại lượng vật lý và cảm nhận tương ứng
Frequency Carving EQ	Kỹ thuật khoét tần số	Chủ động cắt nhẹ 2–4 dB trên track nhạc tại dải Formant thoại để câu thoại nổi bật
Gain Staging	Phân bổ độ khuếch đại	Kỹ thuật thiết lập mức Gain danh định chuẩn (-18 đến -20 dBFS RMS) qua từng mắt xích
Harmonics / Overtone	Hài / Bội âm	Bội số nguyên của tần số cơ bản
Harmonics / Overtones	Hài âm / Âm cộng hưởng	Các thành phần sóng sin có tần số là bội số nguyên của tần số cơ bản f_0 (hài chẵn và hài lẻ)
Headroom	Khoảng dự trữ an toàn	Khoảng cách tính bằng dB từ mức làm việc danh định lên tới trần méo cắt đỉnh 0 dBFS
HRTF	Hàm truyền liên quan tới đầu người	Nền tảng âm thanh nhị nhĩ
Inverse Square Law	Định luật bình phương nghịch đảo	Quy luật mức áp suất âm suy giảm 6,02 dB mỗi khi khoảng cách nhân đôi trong trường tự do
ITD / ILD	Chênh lệch thời gian / mức giữa hai tai	Hai cơ chế định vị
Longitudinal Wave	Sóng dọc	Sóng cơ học trong đó các phần tử môi trường dao động song song với phương truyền sóng
Loudness (LUFS / LKFS)	Độ to cảm nhận chuẩn hóa	Đơn vị đo lường độ to theo chuẩn ITU-R BS.1770 tích hợp mạch lọc trọng số K và Gating
Loudness Range (LRA)	Dải độ to	Chỉ số mô tả độ động chênh lệch giữa đoạn to và đoạn nhỏ của chương trình (đơn vị LU)
Noise Criteria (NC)	Đường cong tiêu chí tiếng ồn	Tiêu chuẩn đánh giá mức độ ồn nền của phòng thu âm và trường quay chuyên nghiệp
Noise floor	Sàn nhiễu	Mức nhiễu cố hữu của hệ thống
Nyquist sampling theorem	Định lý lấy mẫu Nyquist	Tần số lấy mẫu phải lớn hơn ít nhất 2 lần tần số cực đại

Thuật ngữ tiếng Anh (English Term)	Thuật ngữ tiếng Việt quy chuẩn	Định nghĩa chuyên môn & Ngữ cảnh sử dụng
Phase & Phase cancellation	Góc pha & Triệt tiêu pha	Vị trí trong chu kỳ dao động; triệt tiêu khi lệch pha 180°
Phase (φ) / Polarity (Ø)	Góc pha / Cực tính	Hai khái niệm kỹ thuật khác nhau: góc pha phụ thuộc tần số; đảo cực tác động đồng đều mọi tần số
Phon / Sone	Phon / Sone	Hai đơn vị đo lường độ to cảm nhận chủ quan của tai người
Precedence Effect (Haas Effect)	Hiệu ứng đến trước (Hiệu ứng Haas)	Não hợp nhất các âm đến trong vòng 50 ms theo hướng của âm thanh chạm tai sớm nhất
Proximity Effect	Hiệu ứng cận âm	Hiện tượng tăng vọt dải trầm khi đặt micro định hướng ở cự ly cực gần nguồn phát
Reverberation (RT60)	Thời gian âm vang RT60	Thời gian cần thiết để năng lượng âm vang suy giảm 60 dB sau khi ngắt nguồn âm
Reverberation Time (RT60)	Thời gian âm vang	Thời gian giảm 60 dB
Room Modes	Chế độ cộng hưởng phòng	Các tần số cộng hưởng sóng đứng tạo bụng sóng và nút sóng trong phòng kín
Schroeder Frequency	Tần số Schroeder	Ranh giới phân định chế độ xử lý sóng đứng rời rạc và chế độ âm học hình học thống kê
Sibilance	Âm xuýt	Vùng 5–9 kHz, xử lý bằng de-esser
Sound Blanket	Chăn tiêu âm hiện trường	Phụ kiện dạ nỉ chuyên dụng trải sàn/vách triệt tiêu dội âm tạm thời trên phim trường
Sound Pressure Level (SPL)	Mức áp suất âm	Tham chiếu 20 µPa
Speech Intelligibility	Độ rõ nét của lời thoại	Chỉ số đánh giá khả năng nghe và thấu hiểu chính xác nội dung câu từ của người nói
Standing wave / Room mode	Sóng đứng / Mode phòng	Cộng hưởng vùng trầm trong phòng kín
Thuật ngữ tiếng Anh	Tiếng Việt	Ghi chú
Timbre	Âm sắc	Màu sắc âm thanh được quyết định bởi tỷ lệ chuỗi hài Fourier và đường bao biên độ ADSR
Timbre & Harmonic spectrum	Âm sắc & Chuỗi phổ hài	Dấu vân tay âm thanh tạo bởi tỷ lệ năng lượng các họa âm Fourier
True Peak (dBTP)	Mức đỉnh thực liên mẫu	Phép đo đỉnh sóng có nội suy Oversampling ngăn ngừa méo khi nén định dạng phát hành
Wavelength	Bước sóng	$\lambda = c/f$

TÀI LIỆU THAM KHẢO CHƯƠNG 1

[1] International Telecommunication Union (ITU-R). (2023). Recommendation ITU-R BS.1770-5: Algorithms to measure audio programme loudness and true-peak audio level. Geneva: ITU.

[2] European Broadcasting Union (EBU). (2020). EBU Recommendation R128: Loudness normalisation and permitted maximum level of audio signals. Geneva: EBU Technical Committee.

[3] International Organization for Standardization (ISO). (2023). ISO 226:2023: Acoustics — Normal equal-loudness-level contours. Geneva: ISO.

[4] Advanced Television Systems Committee (ATSC). (2013). ATSC Recommended Practice A/85: Techniques for Establishing and Maintaining Audio Loudness for Digital Television. Washington, D.C.: ATSC.

[5] Audio Engineering Society (AES). (2015). AES TD1004: Recommendation for Loudness of Audio Streaming and Network File Playback. New York: AES.

[6] Society of Motion Picture and Television Engineers (SMPTE). (2010). SMPTE ST 202: Dubbing Stages, Screening Rooms and Indoor Theaters — B-Chain Electroacoustic Response. White Plains, NY: SMPTE.

[7] National Institute for Occupational Safety and Health (NIOSH). (1998). Criteria for a Recommended Standard: Occupational Noise Exposure (DHHS/NIOSH Publication No. 98-126). Cincinnati, OH: NIOSH.

[8] Everest, F. A., & Pohlmann, K. C. (2021). *Master Handbook of Acoustics* (7th ed.). New York: McGraw-Hill Education.

[9] Rossing, T. D., Moore, R. F., & Wheeler, P. A. (2014). *The Science of Sound* (4th ed.). Boston: Pearson.

[10] Toole, F. E. (2018). *Sound Reproduction: The Acoustics and Psychoacoustics of Loudspeakers and Rooms* (3rd ed.). New York: Routledge / Focal Press.

[11] Katz, B. (2014). *Mastering Audio: The Art and the Science* (3rd ed.). Burlington, MA: Focal Press.

[12] Rumsey, F., & McCormick, T. (2021). *Sound and Recording: Applications and Theory* (8th ed.). London: Routledge.

[13] Holman, T. (2010). *Sound for Film and Television* (3rd ed.). Oxford: Focal Press.

[14] Zölzer, U. (2008). *Digital Audio Signal Processing* (2nd ed.). Chichester, UK: John Wiley & Sons.

[15] Roads, C. (1996). *The Computer Music Tutorial*. Cambridge, MA: MIT Press.

[16] Smith, J. O. (2011). *Spectral Audio Signal Processing*. Stanford University: W3K Publishing.

[17] Fastl, H., & Zwicker, E. (2007). *Psychoacoustics: Facts and Models* (3rd ed.). Berlin: Springer-Verlag.

[18] Pohlmann, K. C. (2010). *Principles of Digital Audio* (6th ed.). New York: McGraw-Hill.