聲音的隱藏結構：從振動的弦到聲音的形狀

(互動投影片)

陳宜良

台灣大學數學系

2026 年 2 月 7 日

演講大綱

- 聲音、振動與頻率
- 聲波——振動的幾何結構
- 音色——聲音的指紋
- 耳蝸——作為時頻分析器的耳朵

在演講過程中，我們將多次暫停，進行簡短的互動示範：

聲音、振動與頻率

描述振動的方式

- 三角函數：週期性、頻率、振幅、相位，以及音高與頻率的關係
- 疊加兩個三角函數：干涉、協和、不協和、拍頻（beats）

音高與頻率（聽覺 + 量測）

工具

- 聲音範例（觀察聲音的波形）：https://freesound.org/
- 虛擬鋼琴：https://www.musicca.com/piano
- 線上音調產生器：https://onlinetonegenerator.com/
- Frequency Generator（iOS）：產生單一頻率的聲音
- Sonic Visualiser：https://www.sonicvisualiser.org，用於聲音分析
- Sonic Tools（iOS）：在 App Store 搜尋Sonic Tools SVM，用於聲音分析

活動目標

- 播放一個音符（例如 A4），並讀取其對應的頻率。
- 比較相差一個八度的音符：頻率加倍，但「音名」重複。
- （選擇性）唱出一個音，並觀察其主要頻譜峰值。
- 將兩個具有相同頻率、但相位不同的聲音相加——觀察相消或變得更大聲。
- 觀察拍頻（beat）現象。

傅立葉相量（Fourier Phasor）

- 神奇的歐拉公式 $e^{ix} = \cos x + i \sin x$。
- 看見振動：相量（phasor）軌跡

$$A_1 e^{i\omega_1 t + \phi_1} + A_2 e^{i\omega_2 t + \phi_2}$$

  - 干涉（interference）
  - 協和與不協和（consonance and dissonance）
  - 拍頻（beats）
  - 無理頻率比（irrational ratio）

傅立葉展開與傅立葉轉換

任何函數都可以表示為傅立葉級數嗎?

$$x(t) = \sum_{n=-\infty}^{\infty} c_n e^{int}$$

- 一般函數的傅立葉展開、三角級數、相量（phasor）軌跡
- 傅立葉轉換（Fourier transform)、離散傅立葉轉換（DFT)、短時傅立葉轉換（STFT）

Interactive break: 相量軌跡（Phasor trajectories）

(Interactive)

Tools (open these links)

- **Desmos**（傅立葉旋輪示範）：`https://www.desmos.com/calculator/hmsk9bu106`
- **傅立葉旋輪（Fourier Epicycles）**（Dynamic Math）：`https://dynamicmath.xyz/fourier-epicycles`
- **複數傅立葉級數動畫**：`https://fourier.ajai.dev`

Activity goal

- 觀察兩個具有相同頻率、但初始相位相反的相量所產生的相消現象
- 改變頻率比 f_2/f_1，並觀察複平面中的軌跡：協和與不協和
- 簡單的整數比（例如 2:1, 3:2）$\Rightarrow$ 封閉且穩定的軌道
- 複雜或無理的頻率比 $\Rightarrow$ 非封閉、準週期性地填滿平面
- 任意封閉曲線皆可用傅立葉級數來展開

離散傅立葉轉換（DFT）

令 $\omega := e^{-2\pi i/N}$，並定義矩陣

$$F_N := \left[\omega^{kn}\right]_{k,n=0,N-1} = \begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & \omega & \omega^2 & \cdots & \omega^{N-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{N-1} & \omega^{2(N-1)} & \cdots & \omega^{(N-1)^2} \end{bmatrix}_{N\times N},$$

$$\mathbf{x} = \begin{bmatrix} x_0 \\ x_1 \\ x_2 \\ \vdots \\ x_{N-1} \end{bmatrix}, \qquad \mathbf{X} = \begin{bmatrix} X_0 \\ X_1 \\ X_2 \\ \vdots \\ X_{N-1} \end{bmatrix}.$$

則離散傅立葉轉換及其反轉換可寫為

$$\mathbf{X} = \frac{1}{N} F_N \mathbf{x}, \qquad \mathbf{x} = \overline{F_N}^{\,T} \mathbf{X}.$$

短時傅立葉轉換（STFT）

加窗的 DFT

為了追蹤隨時間變化的頻率成分，我們對短且重疊的時間窗進行離散傅立葉轉換。對於離散訊號 $x[n]$、視窗函數 $w[n]$ 與步長（hop size）R，定義

$$X(m,k) = \sum_{n\in\mathbb{Z}} x[n]\; w[n - mR]\; e^{-i2\pi kn/N}, \qquad k = 0, 1, \ldots, N-1.$$

聲譜圖與時頻取捨

聲譜圖（*spectrogram*）定義為 $|X(m,k)|^2$（或 $|X(m,k)|$）。

- 較大的 N：頻率解析度較佳，但時間解析度較差。
- 較小的 N：時間解析度較佳，但頻率解析度較差。

Interactive break: 傅立葉轉換與聲譜圖（STFT）

(Interactive)

Tools (open these links)

- **SpectroDraw**（在聲譜圖上繪製／觀察聲音）：`https://spectrodraw.com/app/`
- Sonic Tools（手機 App） 以手機進行頻譜與音壓量測，適合快速測試與實地實驗。

Activity goal

- 使用自己的聲音：哼唱一個穩定的音高，再向上／向下滑音，觀察聲譜圖中的脊線（ridge）。
- 拍手或打響指：觀察短時間、寬頻的能量爆發。
- （若可調整）改變視窗長度：親身體驗時頻解析度的取捨關係。

振動的總結

- 振動的基本構成單元——透過歐拉表示法的三角函數
- 一般振動可由傅立葉級數來表示
- 傅立葉轉換與離散傅立葉轉換
- 時頻分析——短時傅立葉轉換（STFT）

聲波——振動的幾何結構

一維波動：弦的振動

- 諧振子（harmonic oscillators）
- 振動的弦
- 特徵模態（eigenmodes）：聲音的形狀
- 色散關係（dispersion relation）：特徵值與頻率的關聯
- 一般解：傅立葉展開方法

互動休息：振動弦的特徵模態

工具（請開啟以下連結）

- Falstad：數學物理（振盪與波動）
  `https://www.falstad.com/mathphysics.html`
- HyperPhysics：駐波／弦的振動
  `http://hyperphysics.phy-astr.gsu.edu/hbase/Waves/string.html`

觀察重點

- 固定端弦上的駐波（standing waves）
- 節點（不動的點）與腹點（振幅最大的點）
- 隨著頻率改變而出現的不同模態形狀

活動目標

- 辨識振動弦的特徵模態（eigenmodes）
- 將模態編號與波長及頻率之間的關係連結起來

即時提問

- Q1：在第 1、2、3 模態中，你分別看到了多少個腹點（antinodes）?
- Q2：節點（nodes）位於哪些位置？弦的兩端是否為節點?
- Q3：若模態編號加倍，波長會發生什麼變化?
- Q4：從模態 n 進入模態 $n+1$ 時，共振頻率是否大致等距?
- Q5：若增加弦的張力，共振頻率會上升還是下降？你能否在實驗前就預測結果?

重點整理

對於一條長度固定的弦，其振動可分解為一系列簡單的特徵模態，其形狀為

$$u_n(x) = \sin\left(\frac{n\pi x}{L}\right), \quad n = 1, 2, \cdots$$

對應的波長為 $\lambda_n = \frac{2L}{n}$，共振頻率則為 $f_n = \frac{c}{\lambda_n}$。

- 諧波泛音（harmonic overtones）$\Rightarrow$ 清楚而穩定的音高
- 特徵模態（eigenmodes）$\Rightarrow$ 傅立葉級數的自然基底
- 共振頻率滿足 $f_n = nf_1$，為基頻的整數倍

二維波動：鼓膜的振動

- 二維波動方程
- 正方形鼓與圓形鼓的特徵模態
- 鼓聲的「形狀」
- 色散關係：特徵值與頻率的對應
- 鼓聲的一般形式：正交展開公式

互動休息：鼓膜的特徵模態

工具（請開啟以下連結）

- Falstad：*Mathphysics*（尋找「矩形膜波動」、「圓形膜波動」）
  `https://falstad.com/mathphysics.html`
- Wolfram Demonstrations：圓形鼓膜的正規模態（*Normal Modes of a Circular Drumhead*）
  `https://demonstrations.wolfram.com/NormalModesOfACircularDrumHead/`

觀察重點

- 從振動形狀「聽見」對應的頻率
- 二維膜上的駐波圖樣
- 節線（nodal lines）：完全不動的位置
- 不同幾何形狀對應不同的模態結構

活動目標

- 辨識二維特徵模態與一維弦模態之間的差異
- 將模態形狀與聲音品質連結起來（音高 vs. 音色）

課堂即時提問

- Q1：節線的形狀是什麼？（圓形？直線？）
- Q2：高階模態頻率是否看起來像最低模態頻率的簡單倍數？
- Q3：比較圓形與正方形鼓膜——模態圖樣有何不同？
- Q4：若只在某一點激發鼓膜，會只產生一個模態，還是同時激發多個模態？
- Q5：為什麼鼓的聲音通常較不具有明確的「音高感」？
- Q6：對於一個狹長的矩形鼓，你預期會聽到什麼樣的頻率？

重點整理

二維特徵模態會形成豐富而複雜的空間圖樣，但其對應的頻率通常是非諧和的（*inharmonic*）。
這種幾何上的複雜性，直接形塑了鼓類聲音的音色。

- 弦樂器 $\Rightarrow$ 幾乎為諧和泛音 $\Rightarrow$ 清楚而穩定的音高
- 鼓類 $\Rightarrow$ 非諧和泛音 $\Rightarrow$ 嘈雜或打擊性的音色

你能「聽出」鼓的形狀嗎?

- 一維情形
- 二維情形：面積、周長等幾何量
- 反例

互動休息：你能聽出鼓的形狀嗎?

工具（請開啟以下連結）

- 維基百科：Hearing the shape of a drum

觀察重點

- 不同的形狀（幾何）$\Rightarrow$ 不同的特徵模態
- 每一個特徵模態都對應一個頻率：所有頻率的集合稱為頻譜
- 關鍵問題：頻譜是否能唯一決定幾何形狀?

活動目標

- 理解頻譜（spectrum）所包含的資訊內容。
- 認識為何不同的幾何形狀可能具有相似的頻譜特徵。
- 為下列敘述做準備：「頻譜只能部分決定幾何形狀。」

課堂即時提問

- Q1：若兩個鼓具有相同的面積，你是否預期它們的最低共振頻率相同？
- Q2：若兩個鼓具有相同的周長，你是否預期它們擁有相同的頻譜？
- Q3：你認為哪些幾何特徵對頻率影響最大？（尺寸？形狀？邊界條件？）
- Q4：若只給你一組共振頻率的清單，你能對鼓的形狀推論出哪些資訊？
- Q5：若要唯一地重建鼓的形狀，還需要哪些額外的資訊？
- Q6：啞鈴形鼓、環形鼓的頻譜會長什麼樣子？

重點整理

鼓的頻譜（即特徵頻率）包含了重要的幾何資訊（例如整體尺度），但一般而言，並不足以唯一決定其形狀。
換言之：**頻譜只能部分決定幾何結構**。

- 在一維情形（弦），頻譜對幾何形狀具有很強的約束力。
- 在二維情形（鼓），不同的形狀可能具有相同的頻譜（確實存在反例）。
- 這說明了「聽覺」雖然強大，但並非完美的幾何偵測器。

房間中的振動

- 三維波動方程
- 邊界條件：反射、吸收
- 特徵模態——駐波
- 特徵頻率

互動休息：房間聲學

工具（請開啟以下連結）

- AMROC 房間模態計算器（矩形房間）
  `https://amcoustics.com/tools/amroc`
- Hunecke 房間特徵模態計算器
  `https://www.hunecke.de/en/calculators/room-eigenmodes.html`

觀察重點

- 房間和弦、鼓一樣，也具有特徵模態（駐波共振）。
- 模態高度依賴房間尺寸（長、寬、高）。
- 低頻模態最為稀疏，且最容易造成「轟鳴」（boomy）的聲學問題。

活動目標

- 觀察房間幾何如何決定共振頻率。
- 判斷哪一個空間尺度主導最低頻（低音）共振。
- 將特徵模態與實際聆聽經驗連結（轟鳴區／死點）。

即時提問

- Q1：輸入一個典型房間尺寸（例如 $4 \times 3 \times 2.7$ 公尺），最低的模態頻率是多少?
- Q2：若增加房間長度 L，最低模態頻率會升高還是降低？為什麼?
- Q3：為什麼聲學問題通常在低頻（低音）最為明顯?
- Q4：如果你坐在某一模態的壓力節點位置，在該頻率你會聽到什麼?
- Q5：即使使用相同的喇叭與音樂，為什麼不同形狀的房間聽起來差異很大?
- Q6：比較狹長房間、雙倍房間與單一房間：若它們具有相同體積，其共振頻率有何不同?

重點整理

房間可視為一個三維共振腔：聲場可分解為由幾何形狀與邊界條件所決定的 **駐波特徵模態**。對於長方體房間，其模態頻率近似為

$$f_{n_x,n_y,n_z} = \frac{c}{2}\sqrt{\left(\frac{n_x}{L_x}\right)^2 + \left(\frac{n_y}{L_y}\right)^2 + \left(\frac{n_z}{L_z}\right)^2}, \qquad n_x, n_y, n_z \in \mathbb{N} \cup \{0\}.$$

- 改變房間尺寸會改變（平移）模態頻率分佈。
- 模態會產生熱點（反節）與死點（節點）。
- 聲學處理（吸音、擴散、低頻陷阱）的目的在於控制反射，並抑制過強的共振，特別是在低頻區域。

聲波的總結

- 在封閉區域中，拉普拉斯算子的特徵模態 是聲波傳播的基本構成單元。
- 相應的特徵值透過色散關係與頻率相連。
- 特徵值集合（頻譜）能夠反映部分幾何資訊，但無法完全決定其底層幾何形狀。

音色——聲音的指紋

音色

- 為什麼不同樂器聽起來不一樣?
- 樂器音色的差異
- 男聲與女聲的音色差異（聲譜圖與共振峰）
- 房間的音色（RT60 ／殘響）

長笛的音色

- **聲源–共振器觀點**（source–resonator viewpoint）
- 氣流噴射激發（jet excitation）與聲源頻譜
- 長笛作為一維共振器的模型，以及其傳遞函數
- 音孔（tone holes）可視為阻抗邊界條件

長笛作為一維共振器（空氣柱模型）

波導模型（線性聲學）

以一維（含耗散）的波動方程來描述管腔內壓力 $p(x,t)$：

$$\frac{1}{c^2}\partial_{tt}p - \partial_{xx}p + \frac{\eta}{c^2}\partial_t p = f(x,t), \qquad x \in (0, L).$$

駐波模態與諧波序列

對於（近似）開口–開口管道，其本徵模態給出

$$f_n = nf_1, \qquad f_1 = \frac{c}{2L}, \qquad n = 1, 2, 3, \ldots$$

耗散效應會引入模態衰減率 γ_n （或在集總耗散模型中以共同的 η 表示）。

聲源–共振器觀點與轉移函數 $H(\omega)$

輸入–輸出關係

設吹口（embouchure）提供一個有效的聲源頻譜 $\widehat{S}(\omega)$。對於一個線性可觀測量（例如某一點的壓力，或端點的體積流率），其頻譜為

$$\widehat{X}(\omega) = H(\omega)\,\widehat{S}(\omega).$$

模態轉移函數

一個標準的模態表示形式為

$$H(\omega) = \sum_{n\geq 1} \frac{C_n}{\omega_n^2 - \omega^2 + i2\gamma_n\omega},$$

因此 $|H(\omega)|$ 可視為在 $\omega = \omega_n$ 附近由多個共振「隆起（humps）」所疊加而成。

轉移函數與品質因子

轉移函數 $|H(\omega)|$：峰值出現在模態頻率 $\omega_n = 2\pi f_n$ 附近。對於開口–開口的管道（open–open bore），$f_n = nf_1$。峰值寬度由阻尼控制（在弱阻尼情況下，$\Delta\omega_n \approx 2\gamma_n$）。

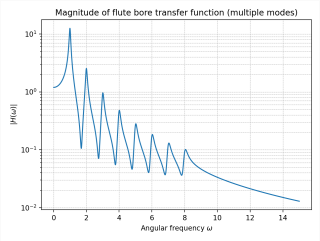


图 1: $|H(\omega)|$。橫軸以基頻正規化，即 ω/ω_1。

品質因子（Quality factor）：

$$Q := \frac{\omega_n}{\Delta\omega_n}$$

品質因子 Q	物理特性	主觀音色感受
大 Q	共振峰狹窄、衰減時間長	明亮、集中、清晰、共鳴感強
小 Q	共振峰寬廣、衰減快速	氣音感重、柔和、擴散

互動中場：長笛的音色

工具（請開啟以下工具）

- Sonic Visualiser （聲譜圖 + 頻譜）：`https://www.sonicvisualiser.org`
- 線上音調產生器 （單一純音 vs. 含泛音）：`https://onlinetonegenerator.com/`
- Sonic Tools（iOS）：在 App Store 搜尋 `Sonic Tools SVM`，用於分析聲音

觀察重點

- 長笛音具有清楚的基頻，且高階泛音相對較少且不強。
- 頻譜包絡線平滑：多數泛音隨頻率上升而快速衰減。
- 在 $X(\omega) = H(\omega)S(\omega)$ 中，管體的轉移函數 H 負責選擇可被放大的共振峰。

活動目標

- 比較長笛在兩個不同音高下的頻譜包絡。
- 辨識共振如何形塑泛音之間的相對強弱。

即時提問（與學生互動）

- Q1：在頻譜中，你能辨認出基頻 f_0 以及泛音 $2f_0, 3f_0, \ldots$ 嗎？
- Q2：隨著頻率升高，泛音振幅衰減的速度如何？
- Q3：當你吹奏較高的音時，泛音之間的比例是否保持不變？哪些量發生了平移？
- Q4：在哪些頻率位置，你看到類似「共振」的峰值（被特別增強的泛音）？
- Q5：哪些部分可以視為聲源 $S(\omega)$，哪些部分像是濾波器 $H(\omega)$？

關鍵重點

長笛的音色主要由管內空氣柱的共振所決定：噴流提供一個相對寬頻的聲源 $S(\omega)$，而管體的轉移函數 $H(\omega)$ 則選擇並加強接近共振頻率的泛音。因此，最終得到的頻譜包絡 $|X(\omega)|$ 通常平滑且相對「純淨」。

互動中場：小提琴的音色

工具（請開啟以下工具）

- Sonic Visualiser：`https://www.sonicvisualiser.org`
- Ave Maria（大提琴 + 鋼琴）

觀察重點

- 小提琴音包含大量強烈的泛音（聲源頻譜相當豐富）。
- 琴身引入額外的共振峰（類似共振峰／ formant 的頻帶），使某些頻率區段被特別增強。
- 頻譜包絡會隨著拉弓位置與力道改變（聲源的變化），也會受到琴身模態的影響（濾波效應）。

關鍵重點

弓弦振動產生一個泛音極為豐富的聲源頻譜 $S(\omega)$。小提琴琴身則作為一個具有多個模態的複雜共振器，形成具有多重峰值的轉移函數 $H(\omega)$。音色正是來自兩者的交互作用：不同泛音被選擇性地放大或抑制，從而塑造出整體的頻譜包絡。

$$A'_k = |H(2\pi k f_0)|\, A_k \quad \Longrightarrow \quad |X(\omega)| \text{ 同時反映弦的泛音結構與琴身的共振特性。}$$

互動中場：人聲的音色

工具（請開啟以下工具）

- Sonic Visualiser：`https://www.sonicvisualiser.org`
- Sonic Tools（iOS）：在 App Store 搜尋Sonic Tools SVM，用於聲音分析
- （選用）線上聲譜圖（快速麥克風示範）：`https://academo.org/demos/spectrum-analyzer/`
- 歌唱時口腔的 MRI 影像

觀察重點

- 人聲由聲帶振動產生基頻及其泛音 $f_0, 2f_0, 3f_0, \ldots$（聲源）。
- 聲道形成共振峰（formants），表現為較寬廣的頻譜峰值（濾波器）。
- 即使音高 f_0 不變，改變母音（/a/、/i/、/u/）仍會使共振峰位置發生位移。

關鍵重點

人聲的音色遵循聲源–濾波器（source–filter）原理：聲帶振動產生具有基頻 f_0 的泛音聲源 $S(\omega)$，而聲道則提供一個具有共振峰的轉移函數 $H(\omega)$，其形狀決定了頻譜包絡。母音的差異，主要來自 $H(\omega)$ 的改變，而非基頻 f_0 的改變。

$$X(\omega) = H(\omega)S(\omega), \qquad \text{共振峰（formants）} \approx |H(\omega)| \text{ 的峰值。}$$

房間也有音色

- 房間是一個三維聲學共振系統。
- 房間模態會選擇性放大某些頻率。
- 牆面吸收與反射決定模態阻尼。

剛性房間的模態

在理想的剛性房間中，聲壓場 $p(\mathbf{x}, t)$ 滿足波動方程式

$$\partial_{tt} p = c^2 \Delta p + s,$$

並在牆面上滿足 Neumann 邊界條件 $\partial_n p = 0$。對於一個長方形房間，利用變數分離法可得到一族駐波本徵模態

$$\phi_{mnl}(x, y, z) = \cos\left(\frac{m\pi x}{L_x}\right) \cos\left(\frac{n\pi y}{L_y}\right) \cos\left(\frac{l\pi z}{L_z}\right), \quad m, n, l \geq 0$$

其對應的本徵頻率為

$$\omega_{mnl} = c\pi \sqrt{\left(\frac{m}{L_x}\right)^2 + \left(\frac{n}{L_y}\right)^2 + \left(\frac{l}{L_z}\right)^2}.$$

房間音色：共振、峰值寬度與衰減

房間轉移函數（模態觀點）

對於每一個具有阻尼的房間模態，

$$H_n(\omega) \;\propto\; \frac{1}{\omega_n^2 - \omega^2 + i2\gamma_n\omega}, \qquad |H(\omega)| \approx \sum_n |H_n(\omega)|.$$

在弱阻尼情況下，會出現類 Lorentz 型的共振峰，其特徵為

$$\Delta\omega_n \approx 2\gamma_n, \qquad Q_n \approx \frac{\omega_n}{2\gamma_n}.$$

對音色的影響

高 Q 模態 $\Rightarrow$ 狹窄的共振峰 + 較長的衰減時間（模態染色明顯，聲音偏「轟鳴」）；
低 Q 模態 $\Rightarrow$ 較寬的共振峰 + 較快的衰減（聲學回應較為中性、平坦）。

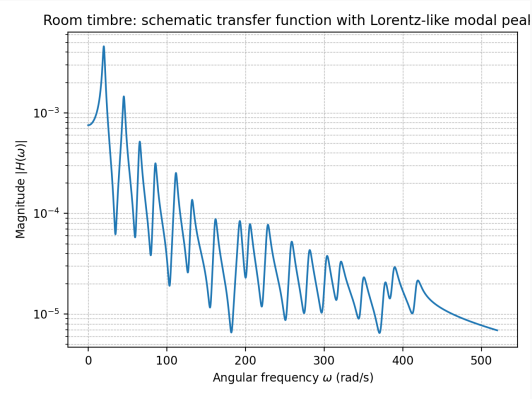


图 2: 共振峰出現在模態頻率 ω_n 附近。峰值寬度由阻尼所控制（在弱阻尼情況下，$\Delta\omega_n \approx 2\gamma_n$）。

混響時間 T_{60}

定義

混響時間 T_{60} 定義為：當聲源停止後，房間中的聲學能量 衰減 $60\,\mathrm{dB}$ 所需的時間，亦即衰減至初始值的 10^{-6}。

與模態阻尼的關係

對於一個具有時間衰減形式

$$q_n(t) \sim e^{-\gamma_n t}$$

的房間模態，其聲學能量隨時間衰減為 $e^{-2\gamma_n t}$。因此，由 T_{60} 的定義可得

$$e^{-2\gamma_n T_{60}} = 10^{-6}, \qquad T_{60}^{(n)} \approx \frac{3\ln 10}{\gamma_n} \approx \frac{6.9}{\gamma_n}.$$

設計上的詮釋

- 較長的 T_{60}：阻尼弱、共振峰狹窄，模態染色明顯。
- 較短的 T_{60}：阻尼強、共振峰寬廣，聲音較為「乾」。
- 房間設計的目標，通常是取得頻率上平衡的 T_{60}，而非單一的通用數值。

混響時間 T_{60}：感知與設計

T_{60}（秒）	主觀聽覺印象
< 0.3	非常乾燥、近似無響室、聲學上「死寂」
0.3–0.6	偏乾、清晰，適合錄音
0.6–1.0	平衡，語音清晰度良好
1.0–2.0	混響感明顯，音樂性豐富
> 2.0	非常「活」、回音感強、空間感寬廣

空間類型	目標 T_{60}（秒）
錄音室	0.2–0.5
教室／演講廳	0.4–0.8
語音演講廳	0.6–1.0
音樂廳（交響樂）	1.8–2.2
歌劇院	1.4–1.8
教堂／大教堂	2.5–4.0

房間設計目標：控制 $Q_n \Longrightarrow T_{60}^n$

- **不要追求單一的 Q**：設計目標在於避免過高的高-Q 共振（尤其是低頻區域），並使各頻段的衰減時間盡量均勻。
- **吸音（Absorption）**：透過能量耗散降低 Q_n（如低頻陷阱、孔隙型吸音材、調諧式吸音器）。
- **擴散（Diffusion）**：重新分配聲能，在不過度阻尼的情況下減少「熱點」與「死區」。
- **反射控制（Reflection control）**：調整早期反射，以提升聲音清晰度與定位感。

$$\text{邊界設計} \Longrightarrow \gamma_n \Longrightarrow \Delta\omega_n \Longrightarrow Q_n \Longrightarrow T_{60}.$$

Interactive break: 房間音色與設計（RT60 ／聲學處理）

(Interactive)

Tools (open these links)

- **AMROC 室內模態計算器** 計算軸向、切向與斜向模態，並視覺化模態密度。
- **RT60 ／混響時間計算器**：`https://app.impulsion-acoustique.fr/reverb`
- **Wayverb（房間脈衝響應模擬）**：`https://reuk.github.io/wayverb/`

Activity goal

- 改變房間尺寸與吸音係數，觀察不同頻段下 RT60 如何隨之變化。
- 嘗試「增加吸音」與「增加擴散」：討論哪些設計有助於提升清晰度，哪些則能保留空間感與寬廣感。

音色總結

- 音色由聲源的頻譜特性，以及共振腔體的轉移函數共同決定。
- 音色是聲音的指紋。

耳蝸——耳朵作為時間–頻率分析器

耳朵作為時間–頻率分析器

- 聽覺系統的解剖結構與機械式濾波
- 耳蝸的結構
- 音調定位（tonotopy）與生物頻譜
- 耳蝸分析作為一種時間–頻率分析

聽覺的解剖結構與機械式濾波

- 聲波進入外耳道，使鼓膜（耳膜）產生振動。
- 中耳聽小骨（錘骨—砧骨—鐙骨）提供阻抗匹配與機械放大。
- 耳蝸是一條螺旋狀、充滿液體的管道，其內部由隔膜分隔成多個腔室（階，scalae）。

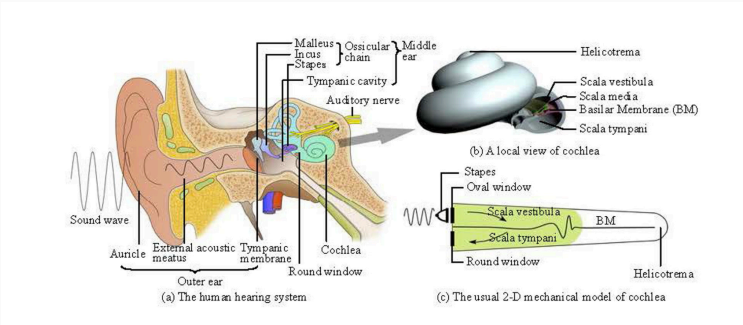
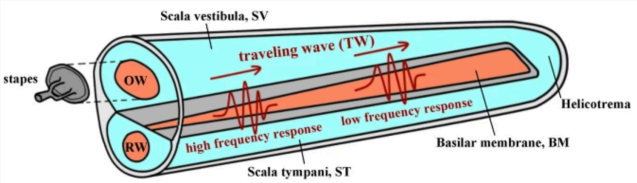


图 **3:** 來源：Ren 等人 [**?**]。

耳蝸的結構

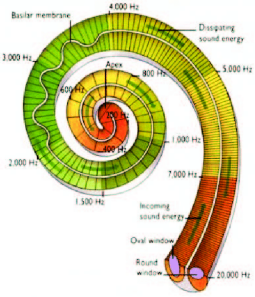
- 鐙骨在卵圓窗處驅動液體運動，在耳蝸內激發波動。
- 基底膜支撐一個行進波，其振幅峰值的位置取決於頻率。
- 機械式濾波發生在毛細胞轉換之前。

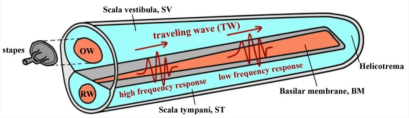


音調定位（Tonotopy）與生物頻譜

- **音調定位（Tonotopy）**：耳蝸沿其長度方向的不同位置，對不同頻率具有最佳響應。
- 高頻在耳蝸基部（較硬、較窄）達到振幅峰值；低頻則在耳蝸頂端（較柔軟、較寬）達到峰值。
- **力學起源**：基底膜在寬度、厚度與剛性上具有系統性的梯度變化，從而產生位置依賴的共振特性。
- 位置—頻率對應關係近似為指數形式：

$$\omega_c(x) = \omega_0 e^{-ax} \quad \Longleftrightarrow \quad \nu_c(x) = \log \omega_0 - ax.$$





八度、音高變化與 $\nu = \log \omega$ 中的平移

八度 = 對數頻率中的加法位移

一個八度對應於頻率加倍：

$$\omega \mapsto 2\omega \quad \Longrightarrow \quad \nu = \log \omega \mapsto \nu + \log 2.$$

因此，在 ν 軸上，等音程對應於等距位移。

泛音結構成為平移的圖樣

對於一個泛音音調 $\omega_k = k\omega^*$，

$$\nu_k = \log \omega^* + \log k,$$

因此，改變音高 ω^*，等同於在 ν 空間中將整個泛音圖樣作平移，並透過 $\nu_c(x) = \log \omega_0 - ax$ 對應為沿耳蝸的位置平移。

重要涵義

音高的改變表現為平移，而音色則編碼在對平移不變的結構之中。

總結：耳蝸（作為時間–頻率分析器）

- 耳蝸對聲音執行一種**機械式的頻率分析**，將頻率映射到基底膜上的空間位置（音調定位，tonotopy）。
- 位置—頻率關係近似為對數形式：

$$\omega_c(x) = \omega_0 e^{-ax}, \qquad \nu = \log \omega,$$

  因此，頻率比例對應於空間上的平移。
- 耳蝸濾波器運作於近似的 **常數-Q（constant-Q）**區域，其對數頻寬為

$$\Delta\nu \approx \frac{1}{Q},$$

  從而提供具有尺度不變性的頻率解析度。
- 音高對應於激發圖樣的**平移**，而音色則編碼在其 **對平移不變的結構**之中。
- 綜合以上特性，耳蝸可視為一個在生物學上高度優化的 **時間–頻率分析器**，與常數-Q 及小波（wavelet）型的訊號表示具有密切關聯。

互動中場：耳蝸如何分析聲音

工具（請開啟以下連結）

- 聽覺機制動畫（Mechanism of Hearing）
  https://youtu.be/m_9SqIQ0BQQ
- Neuroscience Online （耳蝸的音調定位圖）
  https://nba.uth.tmc.edu/neuroscience/m/s2/chapter06.html

觀察重點

- 聲音從耳蝸的基部進入，並以行進波的形式向前傳播。
- 波動的振幅逐漸增大，在某一位置達到峰值後迅速衰減。
- 高頻聲音在基部達到峰值；低頻聲音則在頂端達到峰值。

活動目標

- 將耳蝸濾波理解為一種空間上的頻率分解。
- 將耳蝸的濾波機制與傅立葉分析與 STFT 的概念作比較。

即時提問（與學生互動）

- Q1：對於高頻純音，行進波會在耳蝸的哪個位置達到峰值?
- Q2：對於低頻純音，峰值位置會如何改變?
- Q3：耳蝸中形成的是駐波，還是行進波?
- Q4：為什麼行進波在達到峰值之後會迅速衰減?
- Q5：若同時存在兩個不同頻率，你預期會看到一個峰值，還是兩個?

關鍵重點

耳蝸可視為一個機械式的濾波器組：基底膜上的每一個位置，都對某一個特定的特徵頻率 具有最佳響應，而相鄰位置則對彼此重疊的頻帶作出回應。

- 頻率被映射到空間位置（音調定位，tonotopy）。
- 濾波器具有近似固定的相對頻寬（constant-Q）。
- 這種分析方式更接近小波轉換（*wavelet transform*），而非純粹的傅立葉轉換。

$$\omega_c(x) \sim \omega_0 e^{-ax}, \qquad \Delta\omega \propto \omega_c$$

耳蝸：重點整理

- 耳蝸透過具梯度變化的基底膜，執行一種機械式的頻率分析。
- 音調定位（*tonotopy*）將頻率映射到空間位置，且該對應關係近似為對數刻度。
- 耳蝸濾波器具有近似常數-Q（*constant-Q*）的特性，頻寬約涵蓋一個八度的某一比例。
- 音高（*pitch*）對應於空間上的平移；音色（*timbre*）則對應於圖樣的形狀。

總結

聲音來自振動

- 一般的振動可表示為三角函數的線性組合——傅立葉展開

波動的振動（振動的幾何）

- 拉普拉斯算子的特徵模態——給定幾何中振動的基本構件
- 色散關係——連結特徵值與頻率
- 僅由頻譜本身，無法唯一決定其背後的幾何結構

音色

- 音色由聲源的頻譜特性與共振系統的轉移函數共同決定

耳蝸

- 耳蝸是一個天然、即時運作的傅立葉分析器

結語

從傅立葉模態到生物感知，
聲音揭示了
數學、物理與感知之間的深層統一。

為什麼今天仍然要學習這些主題?

- 這些觀念之所以**古老**，是因為它們**根本而深刻**，而不是因為已經過時。
- **傅立葉分析** → 訊號的表示方式
  - 現代人工智慧同樣在學習「表示」（CNN、音訊模型等）。
  - 卷積在本質上扮演頻率濾波器的角色。
- **特徵模態與頻譜** → 結構與幾何
  - 振動的弦、鼓膜、房間。
  - 頻譜分群、圖神經網路（GNN）。

- **音色與耳蝸** → 多尺度的時間–頻率分析
  - 對數頻率刻度、常數-Q 濾波。
  - 小波轉換與現代音訊特徵。

現代人工智慧並未取代這些觀念——而是在它們之上建構。