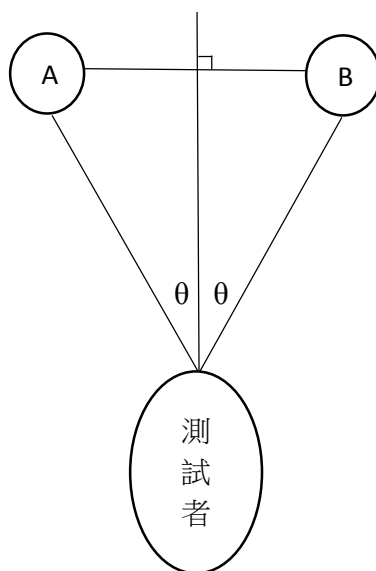


Interaural Phase Difference AND Minimum Audible Angle

1958 年 Mills 做了一個實驗，minum audible angle(MAA)，測試人的耳朵對於聲音的方位辨識。實驗如圖一，以測試者為中心，將聲源 A、B 分別放在測試者左右兩側，讓 A、B 兩點的中點與測試者所形成的直線與直線 AB 垂直。實驗時讓聲源 AB 隨機發出聲音，讓測試者分辨是在左邊或是右邊， θ 為變數，測試發現 θ 在 1° 以內時測試者無法分辨聲音來源為左邊或是右邊。



圖一 MAA 實驗

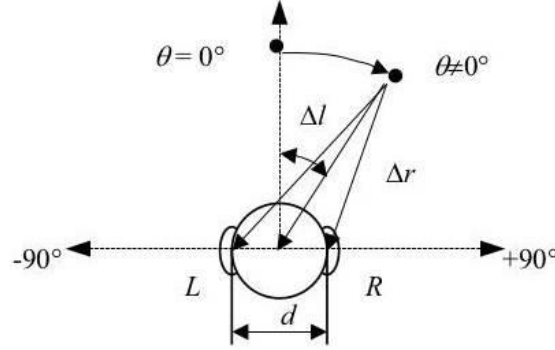
利用 MAA 概念，可以計算出 interaural phase difference(IPD)最大值，IPD 的意義為聲源到左右耳朵的相位角差值，也就是計算差了多少個波長。接下來在介紹如何算出相位差值 IPD 之前，我們先介紹如何計算出聲源到左耳與到右耳的距離差值 Δd ：

圖二為聲源到兩耳的示意圖， Δl 為聲源到左耳的距離， Δr 為聲源到右耳的距離， Δd 為兩者的距離差， d 為左耳到右耳的距離， r 為聲音來源到中心點位置。假設目前方位角為 θ ，則：

$$\Delta r = \sqrt{(r \cos \theta)^2 + (r \sin \theta - \frac{d}{2})^2} \quad (1)$$

$$\Delta l = \sqrt{(r \cos \theta)^2 + (r \sin \theta + \frac{d}{2})^2} \quad (2)$$

$$\Delta d = |\Delta r - \Delta l| \quad (3)$$



圖二 azimuth plane

接下來計算 IPD，IPD(Φ)可表示為頻率 f 的函式：

$$\Phi = \Delta d \times \frac{f}{c} \times 360^\circ \text{ or } \Delta d \times \frac{f}{c} \times 2\pi \text{ radians} \quad (4)$$

其中 c 為聲音在空氣中的傳送速率，約 344m/s，由於 $v = f \times \lambda$ ，故 $\Delta d \times \frac{f}{c}$ 可視為 $\frac{\Delta d}{\lambda}$ ，也就是聲源到兩耳的波程差，最後乘上 2π 或 360° 變成相位角。利用 MAA 概念，當 $\theta \leq 1^\circ$ 時， Φ 計算出的最大值 Φ_{\max} 為 $-3.104 \times 10^{-3} \times f$ ，也就是說，左右耳所接收到特定頻率的波的相位角差值小於 Φ_{\max} 的話，就能夠嵌入資料。

計算左右聲道相位角：

首先先理解什麼是相位角，以下為正弦波表示法：

$$X(t) = A \sin(2\pi f t + \varphi) \quad (5)$$

A 為振幅， f 為頻率， φ 為相位角。由於聲音的資料為時域上的振幅，故我們需要利用傅立葉轉換來得到頻域上的相位，在這邊我們簡單介紹傅立葉概念：傅立葉轉換所得到的數據為傅立葉級數裡的傅立葉系數，而傅立葉級數為無窮多個週期函數來表達原本的函數：

$$F(x) = \sum a_k \sin(kx) + \sum b_k \cos(kx) = \sum C_k e^{ikx} \quad (6)$$

式子(6)中的 C_k 為傅立葉轉換後得到的複數。若將複數 $C_k = N_k(\cos\varphi - i\sin\varphi)$ 代進傅立葉級數裡，則會有有趣的發現，在這邊用簡化的例子來說明：

$$\begin{aligned} F(x) &= \sum N_k (\cos\varphi - i\sin\varphi)(\cos\theta + i\sin\theta) \\ &= \sum N_k [(\cos\theta\cos\varphi + \sin\theta\sin\varphi) + i(\sin\theta\cos\varphi - \cos\theta\sin\varphi)] \\ &= \sum N_k [\cos(\theta - \varphi) + i\sin(\theta - \varphi)] \end{aligned}$$

也就是說，將傅立葉轉換得到的複數以極式表示，此複數的主幅角就是某個頻率的相位角。得到的複數 $a + bi$ ($a, b \in \mathbb{R}$)，利用 \arctan 來計算主幅角：

$$\varphi = \tan^{-1}\left(\frac{b}{a}\right) \quad (7)$$

此複數的振幅，也就是複數平面上的半徑：

$$A = \sqrt{a^2 + b^2} \quad (8)$$

注意這邊的 A 為頻域上的振幅。

接下來就能計算什麼時候能嵌入資料，只要左右聲道相位角差值小於 Φ_{\max} 變能嵌入，在這邊用餘弦來判斷：

$$\cos\{\text{phase}[X_R(f_i)] - \text{phase}[X_L(f_i)]\} < \cos(-3.104 \times 10^{-3} \times f_i) \quad (9)$$

$\text{phase}[X_R(f_i)]$ 為右聲道在頻率為 f_i 時的相位角， $\text{phase}[X_L(f_i)]$ 為左聲道在頻率為 f_i 時的相位角，由於 $\cos(x) = \cos(-x)$ ，不用考慮相位角差值正負，相位角計算方法為式子(7)。

式子(9)即為 psychoacoustical phase threshold，當此頻率滿足式子(9)時，變能夠進行資料隱藏的動作。

參考資料：[IPD](#) [MAA](#)