第二章　理論分析

2.1　本質模組函數（intrinsic mode function，簡稱 IMF）

物理上要定義一個有意義的瞬時頻率（instantaneous frequency）（Copson,1967；Gabor,1946）的必要條件是決定於函數相對於局部零均質（local zero mean）是不是對稱，且有相同數目的過零點（zero crossing）和極值（extrema）。而本質模組函數的定義如下（Huang et al.,1998）：

一、 任何函數的過零點和極值數目相等或兩者最多只差一個。

二、 在數據的任何時間點上，由局部極大值所構成的極大值包絡線（maxima envelope）與局部極小值所構成的極小值包絡線（minima envelope）必須對稱（圖 2-1）。

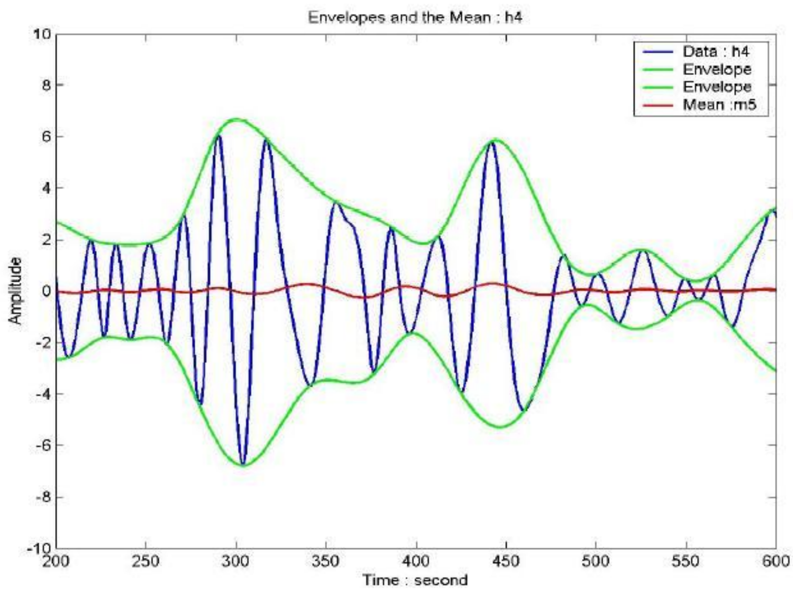

圖 2-1　信號經 cubic spline 連結後所得到之包絡線（Huang, 2003）

上述第一個定義與傳統平穩高斯過程（stationary Gaussian process）中窄頻寬（narrow band）的要求很類似；而第二個定義則將傳統整體性的變化改為局部性的變化。理想上信號局部均值（local mean）應為零值，但對於非穩定性的信號而言需要一個局部的時間尺度（local time scale）來定義一個局部均值，此處的界定是根據採用兩個連續極值之時距（time lapse）。

根據上述所定義出來的本質模組函數表示一個簡單震盪運動，所以它是對應於簡諧函數。每一個 IMF 只包含了一個模態的振動，不會有很複雜的載波（riding wave）。但 IMF 相較於簡諧函數更加一般化，因為它可以同時有調幅和調頻兩種功能（amplitude and frequency modulation）。

2-2　經驗模組分解（empirical mode decomposition，簡稱 EMD）

建立 IMF 是為了滿足 Hilbert-Huang transform 對於瞬時頻率限制條件的前置作業，但絕大部分的震波數據資料都無法符合 IMF 的基本定義，在任何的給定時間之內，都包含了不只一個振動模組，這也是 Hilbert transform 不能對完整的信號提供全盤的頻率內涵的原因。因此必須先將信號數據分解成 IMF，故引入經驗模組分解法來處理非線性及非穩定性的數據。

EMD 能將信號分解成具可調式（adaptive）的 IMF 分量，因為分解的基底函數是從原始信號推導而來的。根據經驗利用信號中特徵時間尺度來定義其振動模組，並依據它來分解信號。如此不但可提供良好的模組解析度，而且能應用到非零均值的資料上，甚至是全無過零點的資料。

經驗模組分解是基於下列幾個假設：

一、 待分析的信號必須至少含有兩個極值，包括一個極大值與一個極小值。

二、 信號特徵時間尺度定義採用兩個連續極值之時距（Dazin,1992）。

三、 如果信號中全無極值只包含反曲點，可以將信號微分一次或多次，將極值找出來，最後的結果可由分量積分得到。

這是用一個有系統的方法來解析出 IMF，又可稱為篩選過程（sifting process），因為在不同階段將不同大小的分量移除。這個分解法分別使用局部極大值和極小值所定義的包絡線，先找出信號中所有的局部極大值，然後利用立方弧線（cubic spline）把它們連接起來當作上圍的極大值包絡線；再找出信號中的局部極小值，同樣使用立方弧線產生下圍的極小值包絡線。上圍和下圍的包絡線應該會包含整個資料。

取極大值包絡線與極小值包絡線的均值包絡線（mean envelope）稱之為 m_1，而原始信號數據 x(t)與 m_1 的差

$$x(t)-m_1=h_1 \qquad (2.1)$$

即為第一個分量，稱之為 h_1。

理論上，h_1 應該是一個合乎 IMF 要求的函數，但原始數據突高突低的情況是很常見的，實際上在處理數據時對極值作迴歸，所以常有越線的數據（overshoots and undershoots），這些數據可用來產生新的極值，並且會將原先的極值移動或放大。即使立方弧線與數據的吻合度極佳，但在斜率上些微隆起的情況也會在將來變成局部的極值，因為當我們在作 $x(t)-m_1=h_1$ 時，已經將直角座標的局部參考零值線移到 m_1，如此便成為曲線座標系統，也就是已經把均值包絡線和 y 軸當成新的座標系統。

在第一輪篩選過後，隆起的值會成為 h_1 的局部極大值，而篩選過程有兩個目的，即為去除載波並將波面（wave profile）變得更對稱。接下來，篩選過程必須重複很多次，方可達到上述目的，因此在第二次篩選過程中，h_1 就當作待處理的數據，故

$$h_1-m_{11}=h_{11} \qquad (2.2)$$

m_{11} 是新的數據組 h_1 的平均值，h_{11} 就是新的數據組與新的平均值的差，經過第二次篩選之後可得到更佳的結果，以此類推可重複

至 k 次，直到 h_{1k} 是 IMF 為止，即

$$h_{1(k-1)} - m_{1k} = h_{1k} \quad (2.3)$$

我們稱 h_{1k} 為數據資料的第一個 IMF 分量，記為

$$h_{1k} = c_1 \quad (2.4)$$

c_1 與 h_1 的差別在於 h_1 仍有可能不完全符合 IMF 的要求，如在兩個過零點之間有多個極值。若篩選過程達到極致，將會抹殺部分具物理意義的振幅擾動，而得到一個純頻率等振幅的信號（如圖 2-2），反而不具真實意義。為了確保 IMF 分量的振幅與頻率變動都能保持足夠的物理意義，篩選過程應該適可而止。設定一些自動停止的準則，如連續三次篩選皆得到相同數目的過零點和極值，便可停止此一模組的篩選過程然後進入下一個模組的篩選過程。圖 2-3 與圖 2-4 即為原始震波數據與其所篩選出來的七個 IMFs 示意圖。

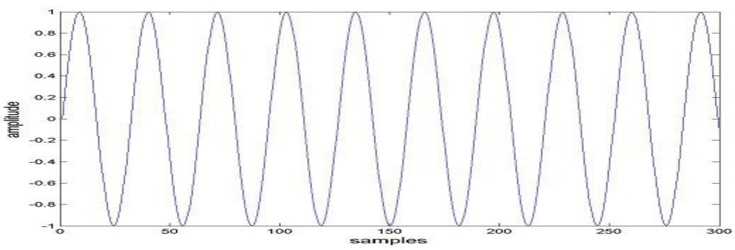

圖 2-2　數據經篩選過多次可能產生不真實的純調頻等振幅信號

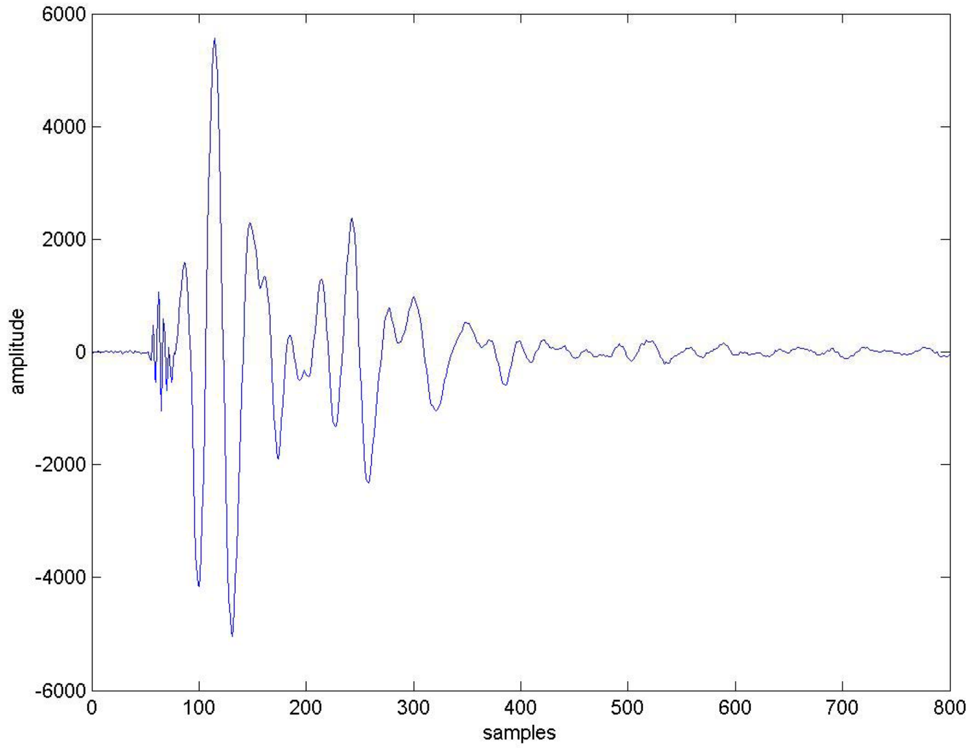

圖 2-3　原始震波信號

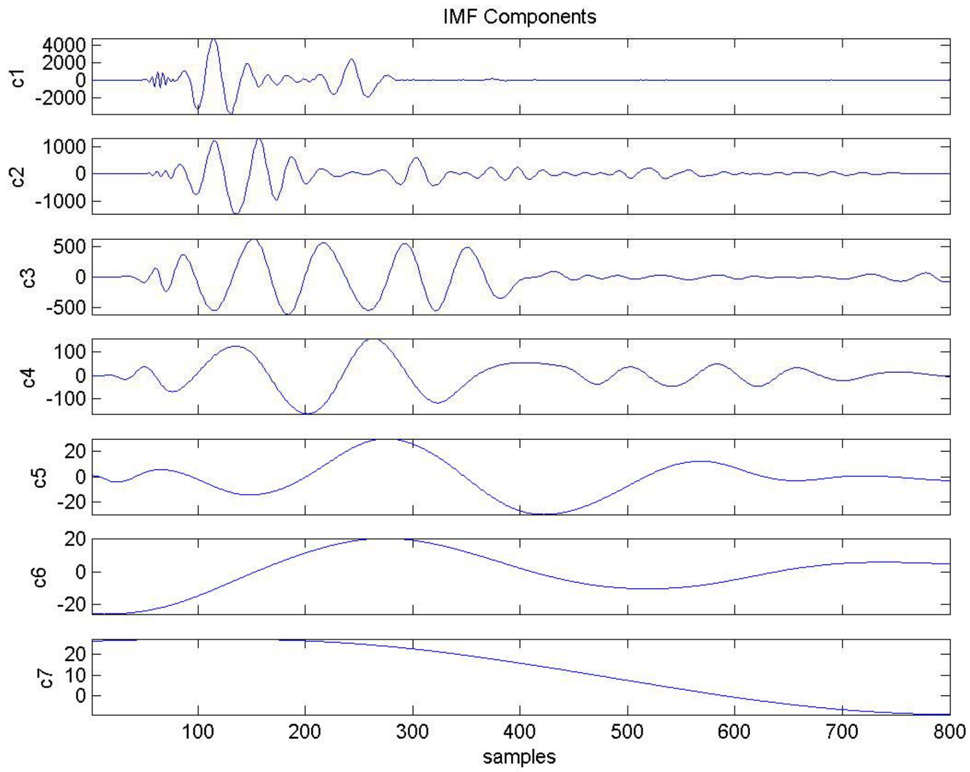

圖 2-4　經重複篩選過程所得七個 IMFs 示意圖

一般而言，c_1 應該包含信號最細微的尺度或最短週期的分量。將 c_1 和原始數據分開，得到

$$x(t) - c_1 = r_1 \qquad (2.5)$$

r_1 為剩餘分量，含有較長週期的分量，可將之視為新的數據，再用上述相同的過程繼續篩選，如此依序可得

$$\begin{aligned} r_1 - c_2 &= r_2 \\ r_2 - c_3 &= r_3 \\ &\vdots \\ r_{n-1} - c_n &= r_n \end{aligned} \qquad (2.6)$$

篩選過程停止的標準可為：

一、 c_n 或 r_n 太小，不具實質意義時。

二、 r_n 變成一單調函數時則無法再提煉出 IMF。

我們可以將原始數據分成 n 個經驗模組（empirical modes）及一個 r_n

$$x(t) - c_1 = r_1，r_1 - c_2 = r_2 \text{ 至 } r_{n-1} - c_n = r_n$$

$$x(t) = c_1 + r_1 = c_1 + c_2 + r_2 = c_1 + c_2 + c_3 + r_3 \cdots\cdots$$

$$x(t) = \sum_{j=1}^{n} c_j + r_n \qquad (2.7)$$

r_n 仍是一組數據，可能是一個均值趨勢（mean trend）或常數。如此，r_n 就不是 EMD 所要的，它是篩選過程中所產生出來的，這同時也提供了一個沒有 DC 困擾的好處。

EMD 分析方法定義出要分析之數據的基本函數集合，數據本身若有改變也會造成基本函數集合的變化，所以這個方法是完全可調（adaptive）的。一旦有了 IMF 分量，便可執行 Hilbert transform

$$x(t,w) = R_e \sum_{j=1}^{n} a_j(t) e^{i\int_i w_j(t')dt'} \tag{2.8}$$

式中 $e^{i\int_i w_j(t')dt'} = e^{i\theta(t)}$ ，$\theta(t)$ 為 Hilbert transform 得到的 phase function，$a_j(t)$ 是振幅，為 Hilbert transform 的實部和虛部平方和的根（$\sqrt{R_e^{\ 2} + I_m^{\ 2}}$）。又因為 r_n 最後為一單調函數，就不能算是頻率的成分，也不計在內了。當然，若 r_n 合格的話，也可以併入在 Hilbert transform 中。