非均勻數據點上之頻譜計算的簡易方法

A Simple Strategy to Evaluate the Frequency Spectrum of a Time Series Data with Non-Uniform Intervals

鄭育能(Yih Nen Jeng)
Department of Aeronautics and Astronautics,National Cheng Kung University
Tainan, Taiwan 70101, Republic of China(**成功大學航太系**)
Email: ynjeng@mail.iaa.ncku.edu.tw
鄭又齊(You-Chi Cheng)
Department of Electrical Engineering
National Taiwan University

ABSTRACT

A simple and fast strategy is proposed to evaluate the frequency spectrum of a string of data whose total number may not be of 2^n and the time intervals between successive data may be non-uniform. The cubic moving least squares method is first employed to separate both the non-sinusoidal and random parts. Next, the Hyunk M3A cubic monotonic interpolation is modified and is employed to convert the remaining data into a data string with uniform time intervals and the total data number is exactly 2^m. Finally, a simple FFT algorithm is employed to provide the spectrum with negligible low frequency error.

Keywords: Cubic Moving Least Squares, FFT, Proper Number of Data, Periodic Condition .

INTRODUCTION

Because of the rapid development of computer hardware and software, the application of a computational fluid dynamical program to unsteady problem becomes a practical issue. During the post processing step, to inspect the frequency spectrum at some typical points is a convenient tool to look into the physical of the flow field. Most computational programs have characters of changing time step size to get merits of both computational stability and computing efficiency. As a consequence, a result of an unsteady CFD program may involve non-uniform time step and have variable step size where number of data points is not exactly equal to some power of 2 or the products of some powers of integers. Consequently, how to employ a fast Fourier transform algorithm to evaluate the frequency may become a difficult issue.

During the developing period of the fast Fourier transform, most people thought that an analogical data string may be exact. On the other hand, the fast Fourier transform is a digital version which can only capture finite number of data. Consequently, most commercial or available fast Fourier transform algorithms have been embedded with the following functions to suppress the aliasing error [1]: side-lobe leakage suppression, adding zeros for circular correlation, and zoom transform etc.. Most of these modifications become trivial, for a result of a computational fluid dynamical or other program, because the output data is principally located at finite points. In other words, it seems that the original FFT algorithm without any modification is more suitable for the post processing than an available FFT program in a commercial software package.

To the author's knowledge, the practical problems of obtaining the frequency spectrum from a result of the computational fluid dynamical program are: (1) how to reasonably removing the numerical error (does the classical statistical methods work?); (2) how to face the problem of non-uniform time steps; and (3) how to treat the problem of data number $\neq 2^m$? In this study, a simple strategy is proposed to solve the last two problems.

ANALYSIS

Cubic Moving Least Squares Method

Consider a set of data, say $(x_i, y_i), i = 0, n$. A moving cubic least squares method defines the error measure function at a point x_k in form of [2]

$$I_k = \sum_{i=0}^{n} e^{-(x_i - x_k)^2/(2s^2)} \left[y_i - f_k(x_i - x_k) \right]^2$$
$$f_k(x) = \sum_{j=0}^{3} a_{kj} x^j \qquad (1)$$

where $e^{-(x_i-x_k)^2/(2s^2)}$ is a Gaussian kernel function with a smoothing factor s. Following the classical least squares method, the minimizing of I_k with respect to parameters a_{kj} resulting a set of linear simultaneous algebraic equations. If the polynomial of $f_k(x)$ takes a constant value, the method becomes a Gaussian smoothing method. If the smoothing factor s is a constant, like the Gaussian smoothing, the FFT algorithm can be employed for the present approach. The required computing count for multiplication and division is $4(n+1)\ln(n+1)$ plus the operating count to evaluate a set of 4 linear equations.

The numerical result of a computational fluid dynamical program often involves non-sinusoidal part, pseudo-sinusoidal and random parts. It is recommended to employ the moving cubic least squares method to isolate the non-sinusoidal and the random parts, with large enough and small enough smoothing factors s, respectively.

Monotonic Cubic Interpolation Method

Consider a Hermite cubic interpolation between points (x_i, x_{i+1})

$$y(x) = c_3(x - x_i)^3 + c_2(x - x_i)^2 + c_1(x - x_i) + c_0,$$
$$c_0 = y_i, c_1 = y_i', \; s_{i+1/2} = \frac{y_{i+1} - y_i}{x_{i+1} - x_i} \qquad (2)$$
$$c_2 = \frac{3s_{i+1/2} - 2y_i' - y_{i+1}'}{x_{i+1} - x_i}, \; c_3 = \frac{y_i' + y_{i+1}' - 2s_{i+1/2}}{(x_{i+1} - x_i)^2},$$

The sufficient monotonic condition of this cubic interpolation is that [3,4]

$$|y'_i|, \; |y'_{i+1}| \le 3s_{i+1/2}, \qquad y'_i \cdot y'_{i+1} \ge 0 \qquad (3)$$

Once the monotonic condition is violated, Fritsch and Carlson [3] proposed to reset y'_i a new value satisfying Eq.(2). In 1992, Hyunk [5] developed several ENO type monotonic cubic interpolants. In this study, his M3A interpolation is employed that gives limiter to the slopes y'_i, y'_{i+1} as

$$y_i' = \mathrm{sgn}(t_i)\min[\frac{1}{2}\left| p_{i-1/2}'(x_i) + p_{i+1/2}'(x_i)\right|, \max(3|s_i|, \frac{3}{2}|t_i|)]$$
$$p_{i-1/2}'(x_i) = s_{i-1/2} + d_{i-1/2}(x_i - x_{i-1}),$$
$$p_{i+1/2}'(x_i) = s_{i+1/2} + d_{i+1/2}(x_i - x_{i+1})$$
$$t_i = \mathrm{minmod}[p_{i-1/2}'(x_i), p_{i+1/2}'(x_i)] \qquad (4)$$
$$d_{i+1/2} = \mathrm{minmod}(d_i, d_{i+1}), \; d_i = \frac{s_{i+1/2} - s_{i-1/2}}{x_{i+1} - x_{i-1}},$$
$$s_i = \mathrm{minmod}[s_{i-1/2}, s_{i+1/2}]$$

At two ends, the Hyunk boundary condition will be employed [5]. As will be discussed later, this cubic interpolation might introduce too much artificial modify-cation.

To the authors' knowledge, the magnitude of spurious oscillation of the cubic spline interpolation is roughly proportional to the ratio of $|y_i / s_{i+1/2}|$ and $|y_{i+1} / s_{i+1/2}|$. For a abrupt discontinuous jump next to a straight line, these ratios might becomes very large. Otherwise, these ratios may be of finite value. Therefore, the desired switching function between the cubic spline interpolation and monotonic cubic interpolation is chosen to be

$$|y'_i|, \; |y'_{i+1}| \le ks_{i+1/2}, \qquad y'_i \cdot y'_{i+1} \ge 0$$
$$k \ge 4 \qquad (5)$$

Simultaneously, in Eq.(4), the slope of monotonic cubic interpolation is modified to

$$y_i' = \mathrm{sgn}(t_i)\min[\frac{1}{2}\left|p_{i-1/2}'(x_i) + p_{i+1/2}'(x_i)\right|, \max(k|s_i|, \frac{k}{2}|t_i|)] \quad (6)$$

In order to reduce error in the limiting case, the cubic interpolation is further degenerated to be a linear equation whenever three successive points are almost collinear, say

$$y_i', y_{i+1}' = 0, \text{ if } |y_{i+1} - 2y_i + y_{i-1}| \le e \text{ or } |y_{i+2} - 2y_{i+1} + y_i| \le e \quad (7)$$

where e is an user specified tolerance. This monotonic interpolation method requires an operation count of multiplication and division in the order of $k \cdot n$ + $(L+3)\cdot m$, where n is number of old data points, $k \approx 30$, m is number of new data points, and L is the count of a searching procedure to allocate an $x \in (x_i^{\mathrm{old}}, x_{i+1}^{\mathrm{old}})$.

Fast Fourier Transform

For the sake of simplicity and to employing the data structure of a computer, it seems convenient to employ the simple fast Fourier transform whose data points are exactly equal to $2^m (= n+1)$ [6]. For a set of data, the Fourier transform pair is

$$y(x) = \frac{a_0}{2} + \frac{1}{2}\sum_{\ell=-\infty}^{\infty}(a_\ell - ib_\ell)e^{i2p\ell f_0 x} = \sum_{\ell=-\infty}^{\infty} a_n e^{i2p\ell f_0 x}$$
$$a_\ell = \frac{1}{T_0}\int_0^{T_0} y(x)e^{-i2p\ell f_0 x}dx \quad (8)$$

where $(0, T_0)$ is the data range, $f_0 = 1/T_0$ is the fundamental frequency. For convenience, in this study, the resulting amplitudes are expressed in terms of their absolute values. From the integration by part formula, it is easy to prove that

$$a_\ell = \frac{T_0}{p\ell}[y'(T_0) - y'(0)] + \frac{2T_0^2}{(2p\ell)^3}\int_0^{T_0} y'''(x)\sin(2p\ell f_0 x)dx$$
$$b_\ell = -\frac{T_0}{2p\ell}[y(T_0) - y(0)] + \frac{2T_0^2}{(2p\ell)^3}[y''(T_0) - y''(0)] \quad (9)$$
$$-\frac{2T_0^2}{(2p\ell)^3}\int_0^{T_0} y'''(x)\cos(2p\ell f_0 x)dx$$

If there are jumps at two ends, the low frequency error will be introduced. In other words, it is would be better to choose data points at two ends with periodic y, y', and y'''s. Since it is not easy to satisfy all of these conditions, it is helpful to choose data at two ends with zero y and y' and y'' being almost periodic. Eq.(9) also shows the fact that a shorter data range gives a smaller low frequency error due to jump at two ends and other artificial modification upon the original data.

Results and Discussions

Consider the original data (shown as thin red line in Fig.1), which is the pressure history around a turbine blade evaluated by a computational fluid dynamics code. The black dotted line is the result smeared by the cubic moving least squares method with $s = 0.05$. A careful inspection upon the difference between the red and black dotted lines (shown as thin blue solid line around the $y = 0$ axis) reveals that the short wave part removed by the moving least squares method is not a well organized composite wave formed from complete sinusoidal waves and noise. In other words, it may involve numerical error and only part of true physics of the flow field for which further studies are necessary. The heavy blue solid line can be considered as the non-sinusoidal part that is obtained by using $s = 0.4$ to smear the black dotted line. The purple solid line is the difference between the black dotted line and heavy blue line. From this example it seems that the cubic moving least squares method is a convenient tool to decouple the random and non-sinusoidal parts from the original data.

For the sake of completeness, a sine wave with wave length $l = 0.5$ (every wave length is resolved by 50 points) is smeared by the Gaussian smoothing (error is shown as thin red line in Fig.2) and cubic moving least squares (as heavy green line) methods, respectively. Those shown in Fig.2 are the maximum error (reduction of the local maximum of the sine wave) generated by two different smoothing methods. It is obvious that, for the Gaussian smoothing, the flatten effect become insignificantly small only if $s < 0.025l$. On the other hand, for the cubic moving least squares method, the flatten effect at the local maximum point is small if $s < 0.1l$. Obviously, the cubic moving least squares has a much better curvature resolving capability than that of the Gaussian smoothing. If a still large smoothing factor s is employed, the sine wave may be flatten to be a straight line. Fig.3 shows the remaining peak value (designed as residue) of both smoothing methods. Their residues are negligible when $s > 0.5l$ for the Gaussian smoothing method and $s > 0.6l$ for the cubic moving least squares method. Their tendencies show that both method's smearing capability are similar. However, their computing time are significantly different.

The effect of the modified Hyunk monotone cubic interpolation can be examined from Fig.4. It seems that, in this example, the present modification does not destroy the monotonic behavior of the original Hyunk cubic monotone interpolation. Figure.5 shows the comparison between the original and modified Hyunk cubic monotone interpolation. The relaxing of the strict monotone condition of Eq.(3) (corresponding to $k = 3$) to be $k = 4$ does change the interpolation shape but still keep monotonic property as shown. From these two test cases, it seems that the present modification partially releases the artificial modifications to some extend. Except for the strange distribution with a linear segment followed by a large jumping, it is generally recommended to employ $k \geq 20$ for most smooth problems.

Now consider the result of applying an FFT algorithm to the dotted line of Fig.1 where the non-periodic condition obvious introduces a large low frequency error as shown in Fig.6. If the linear trend removal method is applied to the dotted line's data, such that the original data is subtracted by a data located on a straight line connecting the initial and final points. The low frequency error is still presented as shown in Fig.7.

In Fig.1, the purple solid line is the result of rearranging the long wave part of Fig.1 (black dotted line subtract the blue solid line) via the modified Hyunk cubic monotonic interpolation [5]. In order to preserve most of the physical characters, at least two new points must be put in every segment between two successive original data points. To reduce low frequency error, data at two ends are truncated so that $y = 0$ at two ends (the resulting T_0 is modified accordingly). The resulting spectrum distribution is shown in Fig.8 which involves small low frequency errors. In order to compare the effect of different Gaussian kernel factor s for the non-sinusoidal part, Fig.9 shows the result of employing $s = 0.6$ to smear the black dotted line of Fig.1. The heavy solid line shows that the non-sinusoidal part estimated by a larger s is more straighten than that with a smaller s. The resulting spectrum distribution is shown in Fig.10. A careful comparison between Figs.8 and 10 reveals that their spectrum distributions are not much different from each other, except that their amplitude magnitudes differ from each other in the order of 10%. For a still larger s, the dominate frequency's amplitude increases about 5% with respect to that of Fig.10. Since most problem does not have a reference to identify the non-sinusoidal part, it is recommended to employ a $s \geq 0.6\, l_{max}$ (which is corresponding to $s = 0.6$ of Fig.10), where l_{max} is the largest wave length estimated by the first dominate frequency.

Figure 11 is the data of vertical displacement at the central point of a steel specimen excited by a hammer. The green line is the original data while the dotted red line is the rearranged data using the modified Hyunk monotonic interpolation. The resulting frequency spectrum is shown in Fig.12 that involves insignificant low frequency error.

Finally, for the sake of completeness, the procedure to employ the present strategy is listed below.

1. Choose the desired data range. For the sake of keeping accurate data evaluation, additional data at two ends are necessary. It would be better to add at least 2 to 3

wave lengths of shortest waves of interesting.

2. Perform the cubic moving least squares method to separate the random and non-sinusoidal part with suitable smoothing factor s 's, respectively (iteration may be necessary).
3. Use the monotonic interpolation method to redistribute the remain data string at proper data points.
4. Find all zero points in the region outside the domain of interest via a simple iteration procedure. From these zero points, choose two end points so that the resulting data range covers the domain of interest. In order to make the error as small as possible, it is necessary to make sure that the zero crossing trends at two ends must be the same, say $\left(y_{l+1}^{\text{left}}-y_l^{\text{left}}\right)\left(y_{r+1}^{\text{right}}-y_r^{\text{right}}\right)>0$, where x_l^{left} and x_r^{right} are those points with $y\approx 0$. Moreover, differences between the first and second order derivatives at two ends must be kept as small as possible.
5. Distribute new data points with the number of new data points $=2^m$ and make sure that at least 2 new points are located in every old data segment.
6. Use a simple FFT algorithm to evaluate the spectrum.

CONCLUSIONS

A simple and complete strategy to reduce the low frequency error and to employ a simple FFT algorithm without any modification is developed. Numerical examples show the robustness of the procedure.

ACKNOWLEDGEMENT

This work is supported by the National Science Council of Taiwan under the grant number NSC-91 -2212-E006-098.

REFERENCES

1. Bendat, J. S, and Piersol, A. G., "Random Data Analysis and Measurement Procedures," 3rd ed., John Wiley & Sons, New York, 2000, Chapters 10 & 11, pp.349-456.
2. Jeng, Y. N., "The Moving Least Squares and Least p-Power Methods for Random Data," The 7-th National Computational Fluid Dynamics Conference, P-9 to P-14, Aug. 2000.
3. Fritsch, F. N. and R. E. Carlson, "Monotone Piecewise Cubic Interpolation," SIAM Numer. Analy. Vol.17, pp.238-246, 1980.
4. de Boor and B. Swartz, "Piecewise Monotone Interpolation," J. Approx. Theory, vol.21, pp.411- 416, 1977.
5. Hyunk, H. T., " Accurate Monotone Cubic Interpolation," SIAM J. Number. Anal. vol.30, no.1, pp57-100, Feb.1993.
6. Brigham, E. O., "The Fast Fourier Transform," Prentice-Hall Inc. Englewood Cliffs, N. J., 1974, pp.164.

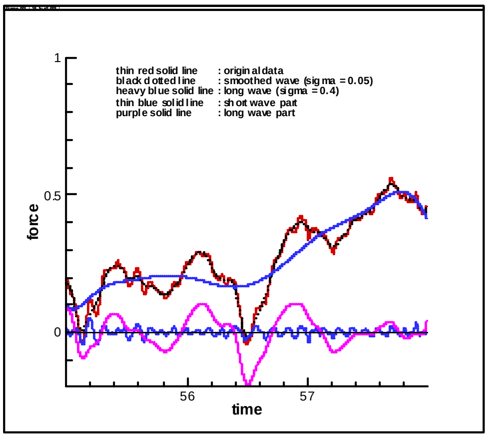

Fig.1 Different parts of data for the pressure distribution around a blade tip: thin red solid line is the original data; black dotted line is the result of employing the cubic moving least squares to smear the local oscillation (irregular short waves); thin blue solid line is the short wave part; heavy blue solid line is the long wave smeared from the black dotted line $s=0.4$; and the purple solid line is the difference between the green dotted line and blue heavy solid line that is the long wave part of the original data.

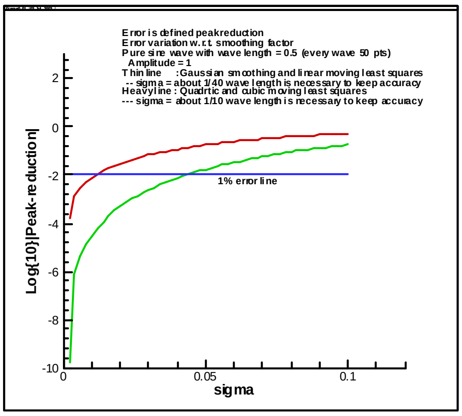

Fig.2 The maximum error at the peak of sine wave with wave length $l=0.5$ (50 points per every wave length) upon smearing of the cubic moving least squares (heavy green line) and the Gaussian smoothing (thin red line).

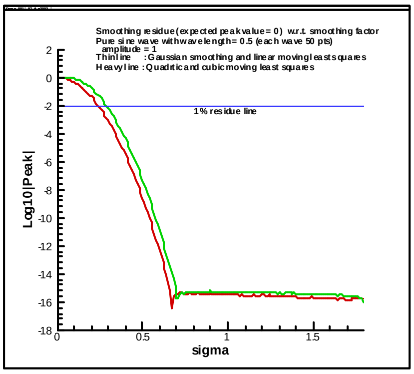

Fig.3 The residue variation with respect to the smoothing factor s : thin red line is result of employing the Gaussian smoothing method and heavy green line is that employing the cubic moving least squares method.

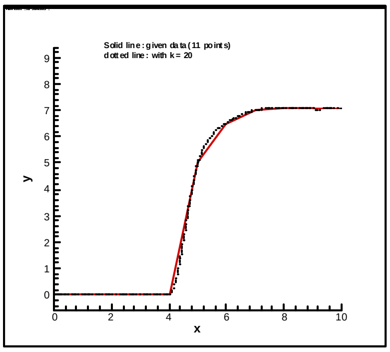

Fig.4 The modification of Eqs.(6,7) does not significantly change the monotone character of the Hyunk monotone interpolation.

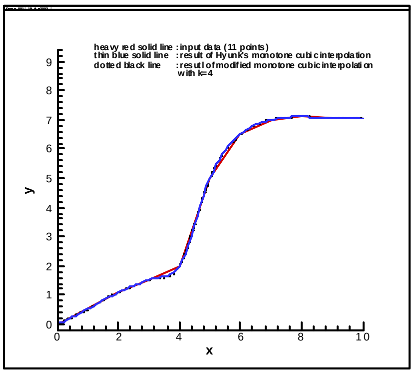

Fig.5 Comparison between results of the original Hyunk monotone cubic interpolation $k=3$ (thin blue line) and the present modification with parameter $k=4$ (dotted line).

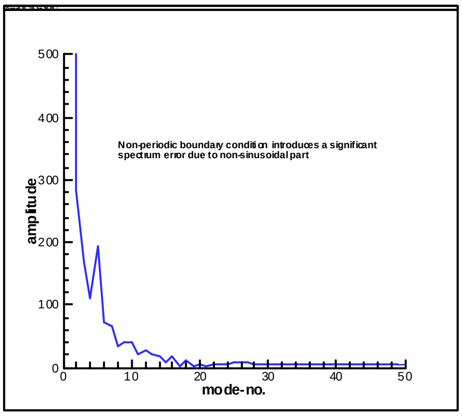

Fig.6 The frequency spectrum of the dotted data of Fig.1, where the large non-periodic boundary condition introduces large errors in the low frequency range.

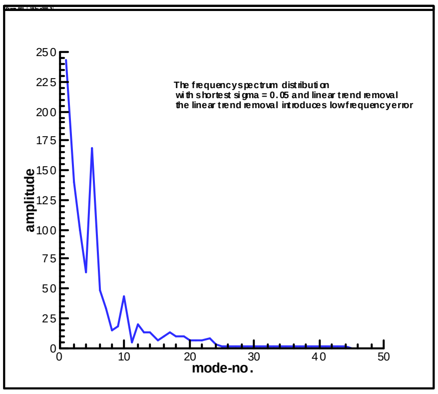

Fig.7 Resulting spectrum of treating the dotted data of Fig.1 via the linear trend removal, the linear trend removal still introduce a significant modification over the low frequency range.

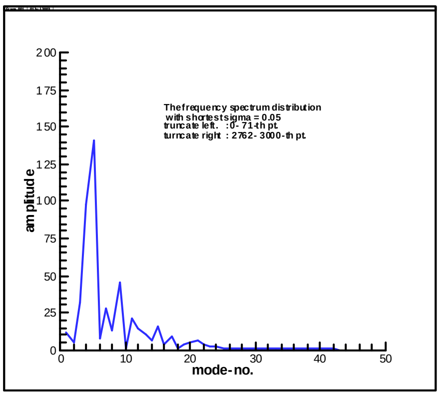

Fig.8 The frequency spectrum of the purple solid line of Fig.1 whose data is truncated at two ends.

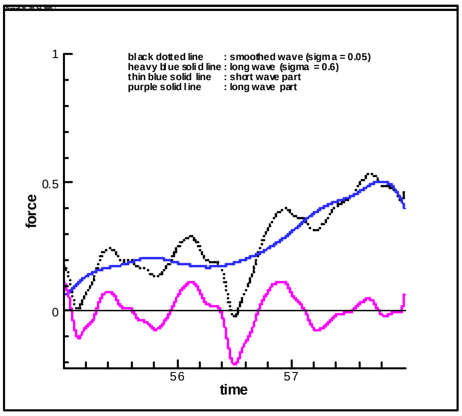

Fig.9 Data for the pressure distribution around a blade tip, heavy blue solid line is the long wave smeared from the black dotted line with $s = 0.6$, other lines are the same as that of Fig.1.

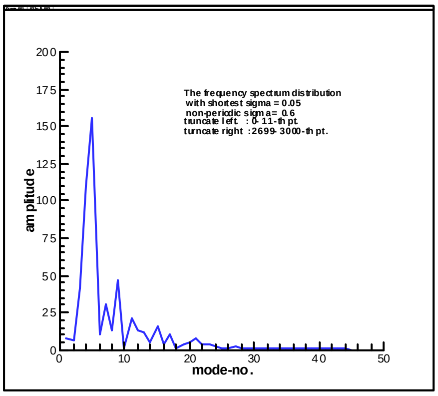

Fig.10 The frequency spectrum of the purple solid line of Fig.9 whose data is truncated at two ends.

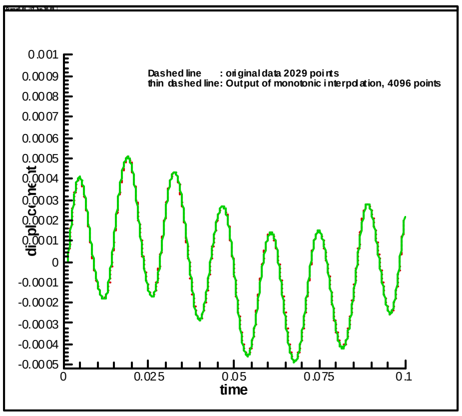

Fig.11 The original data coincides with the rearranged data.

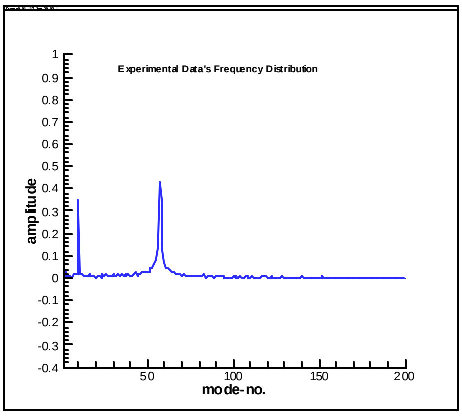

Fig.12 The frequency spectrum with negligible low frequency error.

非均勻數據點上之頻譜計算的簡易方法

鄭育能
成功大學航太系 教授
Email: ynjeng@mail.iaa.ncku.edu.tw
鄭又齊
台大電機系四年級學生

摘　要

本文提出一套簡易之方法，將數據的隨機和非週期函數部份用移動式最小平方誤差法移除，再切除左右兩端點的數據，使之都為為零。不論數據總數等不等於 2^m 及數據間格非均勻，隨後將數據用非震盪式單調內插法重組使數據總點數為 2^m 之均勻數據點。最後應用 FFT 轉換求頻譜。數值測試結果顯示本文方法的結果可以幾乎完全消除低頻誤差。

關鍵詞：FFT 轉換，移動式最小平方誤差法，非震盪式單調內插法，消除低頻誤差。