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

Yih Nen Jeng ** and You-Chi Cheng ‡

**Department of Aeronautics and Astronautics,
National Cheng Kung University
Tainan, Taiwan 701, R.O.C.
‡Department of Electrical Engineering
National Taiwan University
Taipei, Taiwan 106, R.O.C.

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^m 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 Huynh 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.

Key words: Cubic moving least squares, FFT, Proper number of data, Periodic condition.

I. INTRODUCTION

Because of the rapid development of computer hardware and software, the application of a computational fluid dynamical program to unsteady problems has become practical. During the post processing step, inspecting the frequency spectrum at some spatial locations is a convenient tool for looking into the physics of a flow field. Most computational programs have the character 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 steps, possibly steps whose 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 (FFT) algorithm to evaluate the frequency becomes an important task.

During the developing period of the fast Fourier transform, people considered that an analogical data string is exact. On the other hand, an FFT is a digital version which can only capture a finite number of data. Consequently, many available FFT 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 are 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 an FFT algorithm without any modification is more suitable for such a post processing than that with modifications.

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 (do 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 $\neq 2^m$? To the authors' knowledge, the last two problems can be resolved by employing a numerical interpolation algorithm. However, a high order interpolation algorithm may or may not introduce

* Manuscript received, July 29, 2003, final revision, Apr. 16, 2004
** To whom correspondence should be addressed, E-mail: z6208016@email.ncku.edu.tw

spurious oscillation which might introduce significant error to the high frequency part. In this study, a simple and robust strategy is proposed to solve the last two problems.

II. 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]

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

where $e^{-(x_k - x_i)^2/(2\sigma^2)}$ is a Gaussian kernel function with a smoothing factor σ . Following the classical least squares method, the minimizing of I_k with respect to parameters a_{kj} resulting in 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 both the smoothing factor σ and data spacing are constants, 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, pseudo-sinusoidal and random parts. The non-sinusoidal and pseudo-sinusoidal parts are smooth data and reflect the effective data of physics. The random part involves irregular fluctuation with a wavelength in a range of several grid size and is caused by insufficient grid resolution and numerical errors. For an experimental data, the random part is frequently generated by the malfunction of instruments and noise. In order to obtain a relevant physics, people often remove the random part before evaluating the frequency spectrum. Although the smooth non-sinusoidal part is an important part of physics, it contributes a wide band spectrum which introduces low frequency error to the smooth pseudo-sinusoidal part. In this study, the moving cubic least squares method is employed to separate the non-sinusoidal and the random parts, with large enough and small enough smoothing factors σ , respectively.

Monotonic Cubic Interpolation Method

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

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

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

$$|y_i'|,\ |y_{i+1}'| \le 3s_{i+1/2}, \qquad y_i' \cdot y_{i+1}' \ge 0 \tag{3}$$

Once the monotonic condition is violated, Fritsch and Carlson [3] proposed to reset y_i' a new value satisfying Eq.(2). In 1993, Huynh [5] developed several ENO type monotonic cubic interpolation method. In this study, his M3A interpolation is employed that gives limiter to slopes y_i', y_{i+1}' as

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

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

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 an abrupt discontinuous jump next to a straight line, these ratios might become 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

$$|y_i'|,\ |y_{i+1}'| \le ks_{i+1/2},\ y_i' \cdot y_{i+1}' \ge 0,\ k \ge 4 \tag{5}$$

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

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

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

$$y_i' = y_{i+1}', \text{ if } |y_{i+1} - 2y_i + y_{i-1}| \le \varepsilon \quad \text{or } |y_{i+2} - 2y_{i+1} + y_i| \le \varepsilon \tag{7}$$

where ε 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 $x \in (x_i^{\text{old}}, x_{i+1}^{\text{old}})$.

Fast Fourier Transform

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

$$y(x) = \frac{a_0}{2} + \frac{1}{2} \sum_{k=-\infty}^{\infty} (a_k - ib_k) e^{i2\pi k f_0 x} = \sum_{k=-\infty}^{\infty} \alpha_k e^{i2\pi k f_0 x}$$
$$\alpha_k = \frac{1}{T_0} \int_0^{T_0} y(x) e^{i2\pi k f_0 x} dx \tag{8}$$

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

$$a_k = \frac{T_0}{\pi k}[y'(T_0) - y'(0)] + \frac{2T_0^2}{(2\pi k)^3} \int_0^{T_0} y'''(x) \sin(2\pi k f_0 x) dx$$
$$b_k = -\frac{T_0}{2\pi k}[y(T_0) - y(0)] + \frac{2T_0^2}{(2\pi k)^3}[y''(T_0 - y''(0)]$$
$$-\frac{2T_0^2}{(2\pi k)^3} \int_0^{T_0} y'''(x) \cos(2\pi k f_0 x) dx \tag{9}$$

If there are jumps at two ends, the low frequency error will be introduced. In other words, it 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 let y' and y'' be almost periodic. To locate zero y points at two ends, choose two successive points with y_i, y_{i+1} and y_k, y_{k+1} whose values change from positive sign to negative (or from negative to positive), where i denotes a point around the left end and k denotes a point around the right end. Subsequently, a proper interpolation will give the approximate points of zero y. Equations (9) also shows the fact that a shorter data range gives a smaller low frequency error due to jumps of y and its derivatives at two ends and other artificial modification upon the original data.

III. RESULTS AND DISCUSSIONS

In order to examine the effect of moving least squares method, a sine wave with wave length $\lambda = 0.5$ (every wave length is resolved by 50 points) is smeared by the Gaussian smoothing (error is shown as solid line in Fig.1) and cubic moving least squares (as dotted line) methods, respectively. Those shown in Fig.1 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 $\sigma < 0.022\lambda$. On the other hand, for the cubic moving least squares method, the flatten effect at the local maximum point is small if $\sigma < 0.086\lambda$. 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 σ is employed, the sine wave may be flattened to a straight line. Figure 2 shows the remaining peak value (designed as residue) of both smoothing methods. The residues are negligible when $\sigma > 0.48\lambda$ for the Gaussian smoothing method or when $\sigma > 0.58\lambda$ for the cubic moving least squares method. Their tendencies show that both methods' smearing capabilities are similar. However, their computing times are significantly different. From these discussions, it seems that, for a given σ and the cubic moving least squares method, the long waves whose $\lambda \ge 12\sigma$ will be preserved and the short waves with $\lambda \le 1.6\sigma$ will be removed. Therefore, it is an important to search methods for shrinking the transition range.

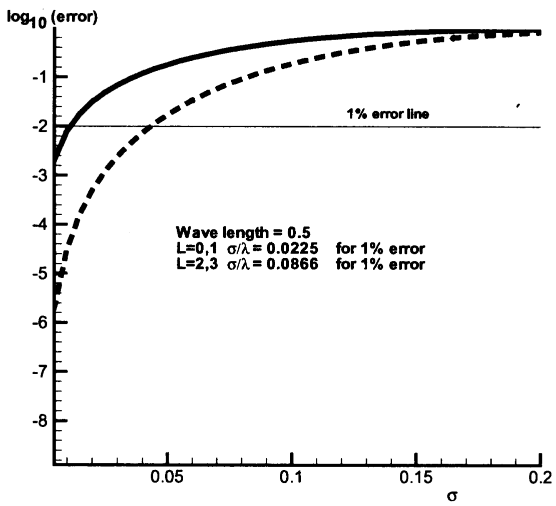


Figure. 1 The maximum error at the peak of sine wave with wave length $\lambda = 0.5$ (50 points per every wave length) upon smearing of the cubic moving least squares (dotted line) and the Gaussian smoothing (solid line).

The effect of the modified Huynh monotone cubic interpolation can be examined from Fig.3. It seems that, in this example, the present modification does not destroy the monotonic behavior of the original Huynh cubic monotone interpolation. Figure.4 shows the comparison between the original and modified Huynh 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 keeps the monotonic property as shown. From these two test cases, it seems that the present modification partially releases the artificial modifications to some extent. 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.

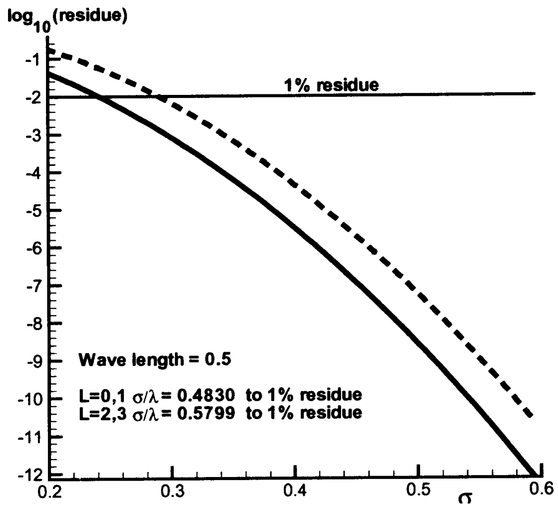


Figure. 2 The residue variation with respect to the smoothing factor σ: heavy solid line is the result of employing the Gaussian smoothing method and the heavy dotted line is the result of employing the cubic moving least squares method.

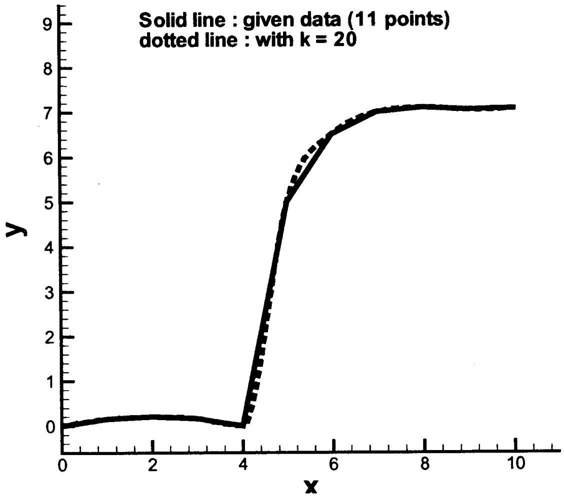


Figure. 3 The modification of Eqs.(6,7) does not significantly change the monotone character of the Huynh monotone interpolation.

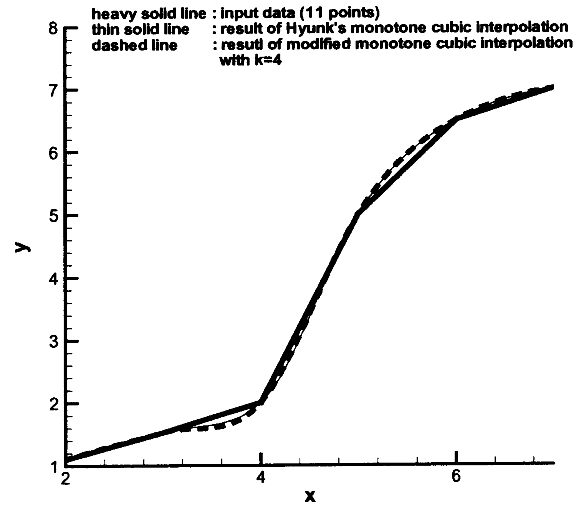
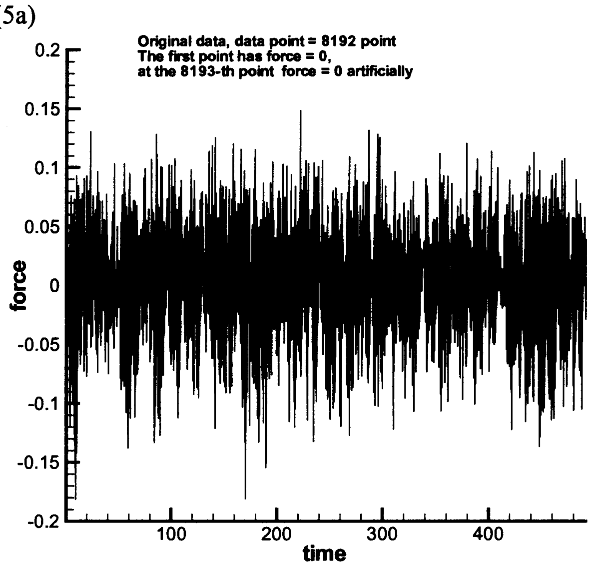


Figure. 4 Comparison between results of the original Huynh monotone cubic interpolation $k=3$ (thin line) and the present modification with parameter $k=4$ (dotted line).

In order to examine the effect upon the frequency spectrum, an experimental pressure data with insufficient sampling rate as shown in Fig.5a and 5b is examined. These data records the turbulent pressure fluctuation (pressure variation with respect to the mean pressure) at the central point of a body's base surface. The axis of log(mode-no) is normalized. In the detailed plot of Fig.5b, the non-smoothness is obvious. At the left end point, the original pressure data is almost equal to the mean pressure within 0.01% so that it is reset to be 0 artificially. For the sake of concentrating on the interpolation effect and to remove the negative effect of non-periodic condition, the subsequent 8191 points are employed, and at the 8192-th point, the pressure is artificially set to zero. Figure.6 shows the spectrum distribution of the pseudo-sinusoidal part on log-log scale. In Fig.7, the result of employing the proposed monotonic cubic interpolation (whose data point increases to be 18392 points) and resulting spectrum is re-scaled to make a clear comparison with respect to that of Fig.6. A careful inspection upon Figs.6 and 7 reveals that, for the data in the range of log(mode no.) <-1.8 of Fig.6, the largest amplitude difference between two figures is less than 1%. Beyond that range, two figures have significant amplitude difference. If the non-smooth data variation is interpreted as the character of random part, this example shows that a numerical interpolation will modify the high frequency part of the spectrum. Therefore, before employing an FFT algorithm, it seems reasonable to remove the random part.

Subsequently, consider the test data (shown as thin solid line in Fig.8), which is the pressure history around a turbine blade evaluated by a computational fluid dynamics code. The line with open circles is the result smeared by the cubic moving least squares method with $\sigma = 0.05$. A careful inspection upon the difference between the thin solid and heavy solid lines (shown as thin solid line around the horizontal 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 study is necessary. The line with open triangles can be considered as the smooth non-sinusoidal part that is obtained by using $\sigma = 0.4$ to smear the heavy dotted line. The line with squares around the horizontal axis is the difference between the open circle line and open triangle line. From this example it seems that the cubic moving least squares method is a convenient tool to decouple the random and smooth non-sinusoidal parts from the original data.

(5a)



(5b)

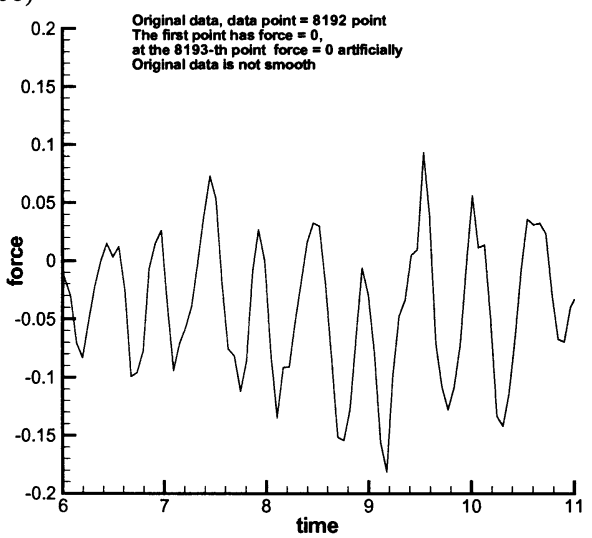


Figure. 5 The pressure fluctuation with respect to the mean pressure at the central point of a bland body's base face: (5a) the overall fluctuation; and (5b) the detailed plot showing the non-smooth behavior.

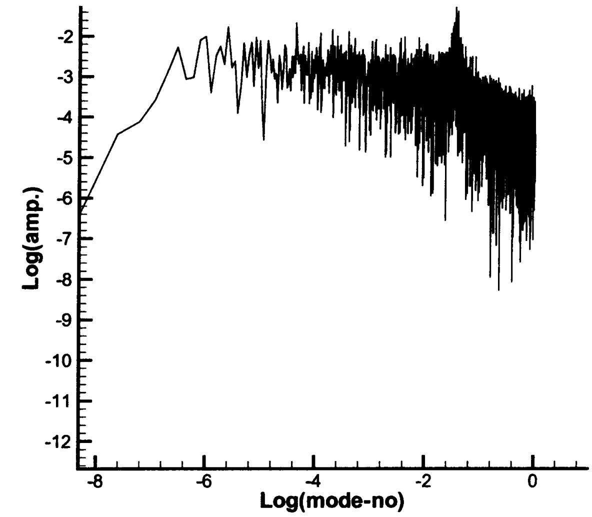


Figure. 6 The frequency spectrum distribution of the original data of Fig.5.

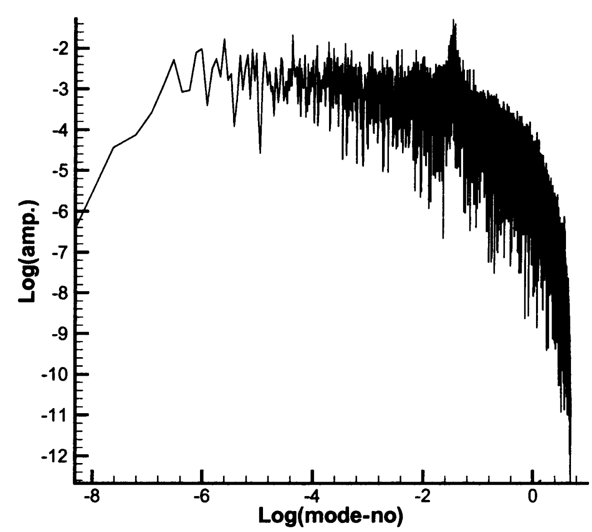


Figure. 7 The frequency spectrum distribution after employing the monotonic cubic interpolation to double the total number of data points.

Now consider the result of applying an FFT algorithm to the dotted line of Fig.8 where the non-periodic condition obviously introduces a large low frequency error as shown in Fig.9. 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.10.

In Fig.8, the line with squares around the horizontal axis coincides with the result of rearranging the long wave part (the open circle line subtracted by the line with triangles) via the modified Huynh 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 properly chosen by the above mentioned method so that force = 0 at two ends (the resulting T_0 is modified accordingly). The resulting spectrum distribution is shown in Fig.11 which involves small low frequency errors. For the sake of comparing the effect of different Gaussian kernel factor σ on the non-sinusoidal part, the line with triangles of Fig.12 shows the result of employing $\sigma = 0.6$ to smear the open circle line.

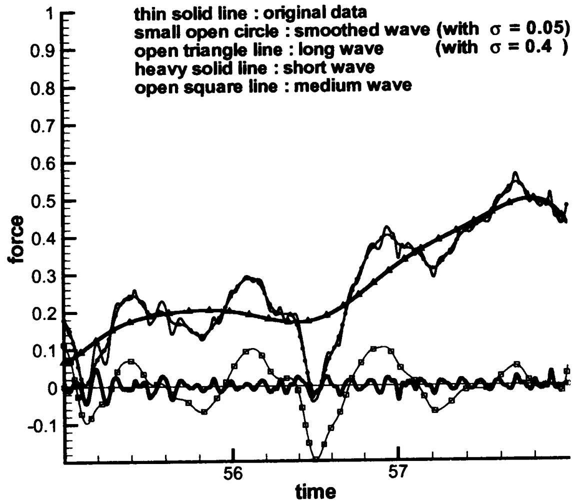


Figure. 8 Different parts of data for the pressure distribution around a blade tip: thin solid line: original data; open circle line: result of smoothed short wave with $\sigma = 0.05$; triangle-line : long wave with $\sigma = 0.4$; square-line around force = 0 ; smoothed long wave minus smoothed short wave; and heavy solid line around force =0: shortest wave.

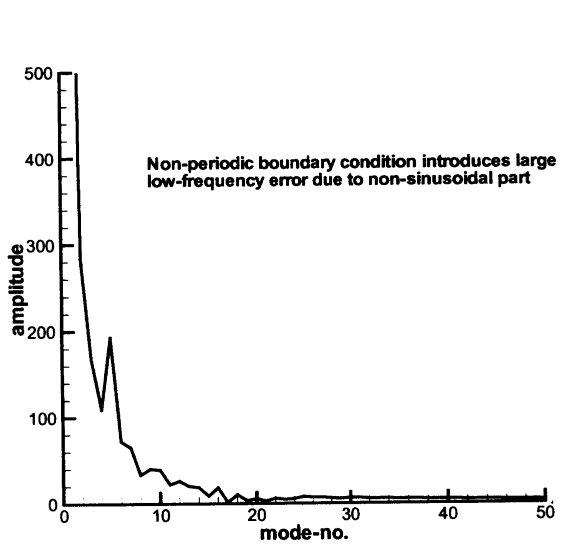


Figure. 9 The frequency spectrum of the line with open circles of Fig.8, where the large non-periodic boundary condition introduces large errors in the low frequency range.

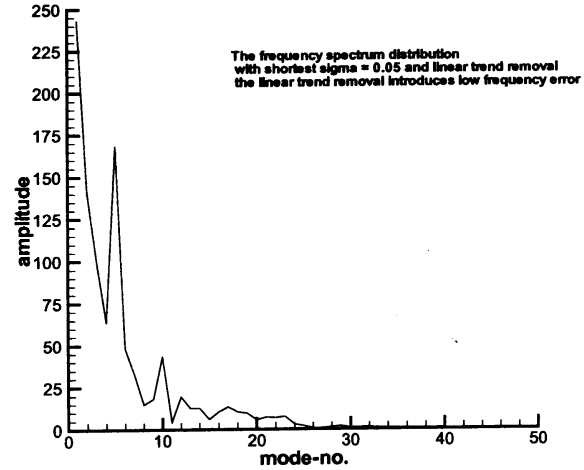


Figure. 10 Resulting spectrum of treating the open circle line of Fig.8 via the linear trend removal, the linear trend removal still introduce a significant modification over the low frequency range.

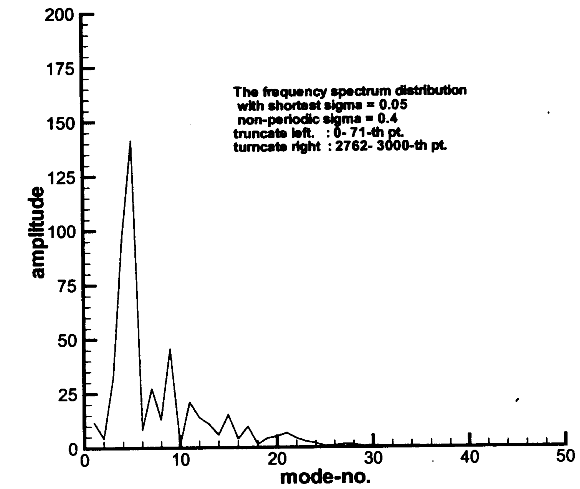


Figure. 11 The frequency spectrum of the line with squares of Fig.8 whose data is truncated at two ends.

The result shows that the non-sinusoidal part estimated by a larger σ is more straighten than that with a smaller σ. The resulting spectrum distribution is shown in Fig.13. A careful comparison between Figs.11 and 13 reveals that their spectrum distributions are not much different from each other, except that their amplitude magnitudes differ from each other in an order of 10%. For a still larger σ, the dominate frequency's amplitude increases about 5% with respect to that of Fig.13. Since most problems do not have a reference to identify the smooth non-sinusoidal part, it is recommended to employ a $\sigma \geq \lambda_{max}$ (which is approximately corresponding to the σ value of Fig.13), where λ_{max} is the largest wave length estimated by the first dominate frequency. On the other hand, to remove the random part, chooses σ to be $2-3\Delta t$ or the approximate wave length of irregular data variation.

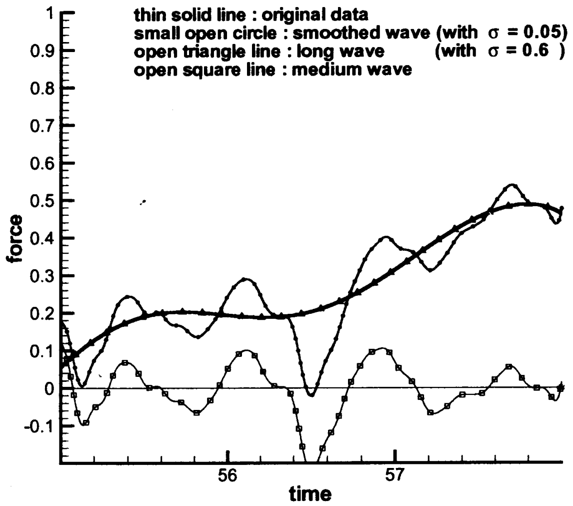


Figure. 12 Different parts of data for the pressure distribution around a blade tip; open circle line: result of smoothed short wave with $\sigma = 0.05$; triangle-line : long wave with $\sigma = 0.6$; and square-line around force =0: smoothed long wave minus smoothed short wave.

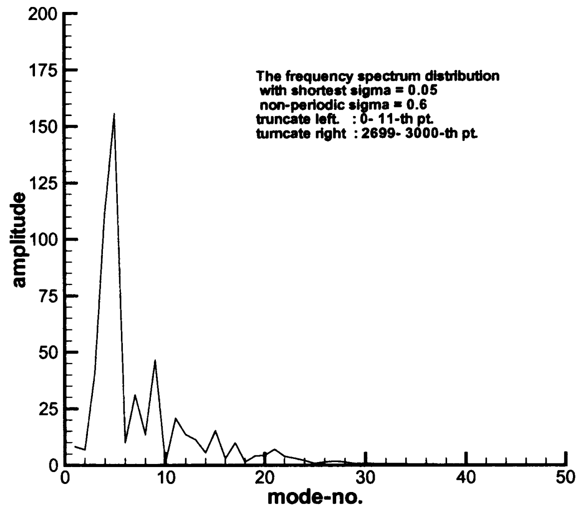


Figure. 13 The frequency spectrum of the line with squares of Fig.12 whose data is truncated at two ends.

As discussed above, in the procedure of decoupling the random and smooth non-sinusoidal part, the magnitude of the Gaussian kernel factor σ plays an important role. Although the present study shows that it should be determined by physical insight of the problem, it still involves certain arbitrariness and non-uniqueness because of the large transition range between the capability of removing short waves and preserving long waves. In other words, extensive further studies are necessary.

Figure 14 is the data of vertical displacement at the central point of a steel specimen excited by a hammer. The solid line is the original data while the open circle symbols are the rearranged data using the modified Huynh monotonic interpolation. The overlapping plot between the original data and interpolated data shows that, if the proposed interpolation method is applied to a smoothed data, the embedded interpolation error is insignificantly small. The resulting frequency spectrum is shown in Fig.15 that involves insignificant low frequency error.

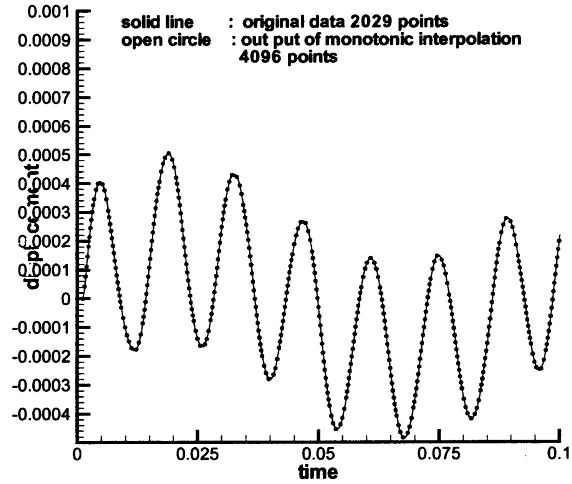


Figure. 14 The original data (solid line) coincides with the rearranged data (open circle symbols).

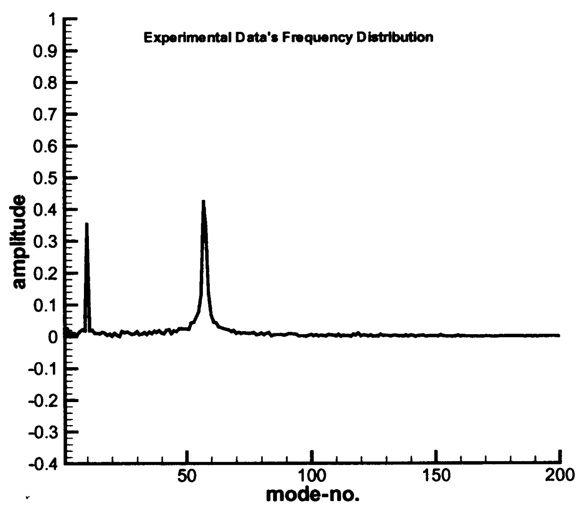


Figure. 15 The frequency spectrum with negligible low frequency error.

Finally, 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 shortest wave lengths of interesting.
2. Perform the cubic moving least squares method to separate the random and smooth non-sinusoidal part with suitable smoothing factor σ 's, respectively (repeated iteration may be necessary).
3. Use the monotonic interpolation method to redistribute the remaining data string at proper data points.
4. Find all zero points in the region outside the domain of interest via a simple inspection 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 $(y_{i+1}^{\text{left}} - y_i^{\text{left}})(y_{k+1}^{\text{right}} - y_k^{\text{right}}) > 0$, where x_i^{left} and x_k^{right} are those points with $y \approx 0$. The points of $y = 0$ at two ends can then be determined by a proper interpolation formula. 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. For a practical application, in order to have consistent spectrum plots, the horizontal axis should be properly scaled so that every wave before and after employing the interpolation algorithm is expressed in the same sine and/or cosine function.

IV. 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] Huynh, 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.