Bull.Ind.Soc.Earth.Tech., Paper No.327, Vol.29, No.4, Dec.1992, pp.53-67

THE EFFECT OF BAND LIMITED INTERPOLATION OF NONUNIFORM SAMPLES ON RECORDS OF ANALOG ACCELEROGRAPHS

Ashok Kumar, S. Basu and Brijesh Chandra
Department of Earthquake Engineering,
University of Roorkee, Roorkee 247667, India.

Abstract

The earthquake records of analog accelerographs which are digitised on semi-automatic digitisers provide uncorrected accelerograms at unequal sample intervals (nonuniform samples). During the processing of such uncorrected accelerograms, most of the researchers usually perform linear interpolation to obtain data at constant sampling interval. However, it is well recognised that linear interpolation introduces frequencies from zero to infinity and the digital data which is always band limited is distorted by linear interpolation in the entire band due to the effect of aliasing. In this paper, an iterative method to recover band limited signal from nonuniform samples is discussed. A sinewave and a set of earthquake records defined at nonuniform samples are recovered at 200 samples per second with the above method. The same sine wave and the earthquake data are also linearly interpolated to get 200 samples per second data and a comparison in frequency domain is done to highlight the limitations of linear interpolation. In the process of this study, several uncorrected accelerograms of past earthquakes are reviewed and their limitations are discussed. In particular, the El Centro accelerogram and the Parkfield accelerogram are studied in detail and comparison of response spectra obtained through linear interpolation and through band limited interpolation is presented.

INTRODUCTION

Interpolation of a given sequence is a requirement which arises frequently while processing digital data. In the field of digital signal processing, such interpolations are usually done in a manner such that the frequency contents of original data do not change. However, in strong motion data processing, researchers mostly perform linear interpolation which distorts the frequency content of digital data. In this paper, the necessity of using band limited interpolation while processing accelerograms is highlighted and a method to perform band limited interpolation for nonuniform samples is suggested.

The trace of most of the past earthquake records of analog accelerographs are digitized on semi-automatic type digitisers the kind of which are still quite commonly employed for this purpose. It is, therefore, obvious that such data cannot be obtained at a constant time interval and these are called nonuniform samples. All the presently available methods of processing of strong motion data convert these nonuniform samples into data at constant time interval generally using linear interpolation. It is well recognized that linear interpolation introduces frequencies from zero to infinity. Digital data which is always band limited, gets distorted by linear interpolation in the entire band upto Nyquist frequency due to effect of aliasing.

Another question of interest is as to what should be the interval at which operator should perform the digitization so that mathematically sound information can be obtained in a desired band.

In this paper, concept of stable sampling set in regard to nonuniform samples is first presented. Based on this, some of the strong motion data of past earthquakes are reviewed and their limitations presented. Further an iterative method is introduced for recovery of signal from nonuniform samples. The signal is recovered at 200 samples per second (SPS) and the iteration is done on the low passed staircase function of the nonuniform samples. Sinewaves as well as earthquake motions defined by nonuniform samples are recovered using the above scheme. The comparison of results of band limited interpolation of nonuniform samples with that of linear interpolation of nonuniform samples is done through Fourier spectra. Limitations of data obtained for El Centro record of May 18, 1940 and Parkfield record of June 27, 1966 are discussed by comparison of response spectra. The efficacy of the band limited interpolation *vis a vis* linear interpolation is examined and highlighted in this paper.

DIGITIZATION OF ANALOG RECORDS AND STABLE SAMPLING SET

Semi-automatic digitisers require judgement of the operator to pickout points on the trace (analog record on photographic paper or film) which is to be digitised. The question is how these points be picked up so that the signal can be recovered within the required frequency band. In this regard the concept of stable sampling set becomes important. Marvasti (1987) and Marvasti *et al.* (1991) defined stable sampling set as follows:

A stable sampling set is a set of digital data that uniquely determines a band limited signal and satisfies the following inequality:

$$\int |x(t)|^2\, dt \leq C \sum_{n=-\infty}^{\infty} |x(t_n)|^2 \qquad (1)$$

where $x(t)$ is the signal and $x(t_n)$ is its n^{th} discrete sample. C is a parameter which represents nonuniformity of the samples.

It has been shown by them that the first condition of stable sampling set *ie.* a set of digital data from which the band limited signal can be determined uniquely, is achieved if the average sampling rate of the digital data is greater than Nyquist rate. However, such sampling set will be stable only when the sufficient condition of Inequality 1 is satisfied. Marvasti *et al.* (1991) have shown that one of the way to satisfy this condition is that all the samples should not be clustered in a finite interval but should be evenly distributed such that

$$|t_n - nT| < \frac{T}{4} \qquad (2)$$

where t_n is the instant of n^{th} nonuniform sample and T is the sampling interval which should be less than or equal to sampling interval corresponding to Nyquist rate.

The same authors have presented an iterative scheme to recover signal from stable sampling set of nonuniform samples (discussed in next section). They have also shown that if in a nonuniform sequence, average sampling rate is higher than Nyquist rate but the Inequality 2 is not satisfied then the convergence of iteration to recover the signal is possible but is slower. Also if Inequality 2 is not satisfied and the average sampling rate is exactly equal to the Nyquist rate the convergence of iteration is not guaranteed. However, if the average sampling rate of the sequence is less than the Nyquist rate then the iteration will not converge. Thus the following can be concluded in relation to the digitization of the trace of an analog accelerogram.

1. If frequencies upto 25 Hz are required to be recovered then average sampling rate should be atleast 50 samples per second (SPS) which is the Nyquist rate. This means that the number of total samples digitized should be at least 50 times the duration of record in seconds. This is a necessary condition to get the desired results. In case the number of digitized samples are less than this value then the maximum frequency that can be recovered from such data will also reduce accordingly.

2. The sampling points should be as evenly distributed as possible and clustering of sampling points around a finite duration should not be done. It will be desirable if the distribution of most of the sampling points satisfy Inequality 2.

3. Picking out peaks and troughs, which is the usual practice in digitizing traces of records of analog accelerographs, is not as important as distributing the sampling points evenly.

4. Although, the condition of digitizing with average sampling rate greater than Nyquist rate is an essential requirement, however, digitization to satisfy Inequality 2 is only a desirable requirement as it increases the speed of convergence. It should be appreciated that it will be impractical for an operator to pick samples during digitization strictly in a manner so as to satisfy Inequality 2.

METHOD OF INTERPOLATION

Once the digitized data of analog trace is obtained, the next operation is to use calibrations of time mark and tilt sensitivity of accelerometer to get uncorrected accelerogram at nonuniform sample intervals. The signal now is required to be recovered from given sequence of nonuniform interval in such a way so as to maintain the same frequency content. The word recovery of signal in this study means getting digital data at comparativily large sampling points at constant interval. In this work, the signal is recovered at 200 SPS.

Marvasti *et al.* (1991) have shown that for nonuniform samples given by

$$x_s(t) = \sum_i x(t_i)\delta(t - t_i) \tag{3}$$

where $\{t_i\}$ is a stable sampling set, the following iterative method shall recover the original band limited signal $x(t)$ from $x_s(t)$

$$x_{k+1}(t) = \lambda PS(x(t)) + (P - \lambda PS)x_k(t) \tag{4}$$

where λ is a convergence constant whose value should be between 0.5 and 1, $x(t)$ is the original signal and $x_k(t)$ is the signal obtained after k^{th} iteration. P is a band limiting operator (a low pass filter) and S is ideal nonuniform sampling operator or a staircase function operator. The band limiting operator P is self adjoint which means $P(x_k(t)) = x_k(t)$. The staircase function operator in this case assumes zero order hold of nonuniform samples and converts the nonuniform samples into a staircase form at a constant interval with 200 SPS. $PSx(t)$ in Equation 4 is low passed signal of staircase nonuniform samples which is known before the start of iterations. Marvasti *et al.* (1991) have also proved that there exist a range of values of λ for which the process of Equation 4 will converge i.e.

$$\lim_{k \to \infty} x_k(t) = x(t) \tag{5}$$

The S Operator

The S operator in iteration of Equation 4 requires construction of staircase from the nonuniform samples. The staircase is obtained by making a sequence at 200 SPS with each sample equal to last available nonuniform sample. Making stair case from a given nonuniform sample really means getting a zero order sample and hold signal. In fact the logic of recovering a signal from nonuniform samples through the staircase function is quite similar to the practice of constructing a band limted analog signal from a given digital data through sample and hold.

The P Operator

The P operator in Equation 4 is called as band limiting operator which is essentially a low pass filter. Butterworth filter is used for performing the low pass operation. The filtering is performed in frequency domain. The filtering operation is performed by first finding FFT of the signal which is to be filtered. The bin frequency interval (Δf) at which the FFT of the signal is determined is given by

$$\Delta f = \frac{SPS}{ndata} \tag{6}$$

where SPS denotes samples per second which is 200 SPS in this case and $ndata$ is the number of data points of the sequence which is lengthened by inserting zeros so that the sequence has exactly 2^j (where j is an integer) data points. For example, if sequence has 3910 data points then $ndata$ is made 4096 (which is 2^{12}) by inserting 186 zeros at the end of the sequence. At the above bin frequency intervals, transfer function $|H(j\omega)|^2$ of Butterworth filter, which is given by the equation below (Lam 1979), is determined from DC to the Nyquist frequency.

$$|H(j\omega)|^2 = \frac{1}{1 - (\frac{\omega}{\omega_c})^{2n}} \tag{7}$$

where ω is the frequency at which the function is determined, ω_c is the cutoff frequency and n is the order of filter. It is quite obvious that cutoff frequency of the lowpass filter should be less than half average sampling rate of the nonuniformly sampled data. The order of the filter (n) determines the the sharpness or the roll off of the filter function. A higher order filter provides steeper cut off or smaller roll off frequency. However, a very large order of filter introduces Gibbs phenomena particularly near cut off frequency. In this study, it has been found that a filter order of 6 provides a filter function which is sufficiently accurate for our work.

Convolution is then performed between the FFT of the signal and the transfer function of the filter. As the transfer function of filter has only real part, convolution in frequency domain means simple multiplication of real and imaginary parts of the signal with the filter function at the corresponding bin frequencies. Inverse FFT is then performed to get the low passed signal in time domain.

Iteration

Figure 1 gives flow chart of the software developed to perform the iteration of Equation 4. The algorithm first converts nonuniform samples into staircase at 200 SPS . FFT of the staircase function is then found and $|H(j\omega)|^2$ of Butterworth filter is used to convolute the signal for low pass filtering in frequency domain. An inverse FFT is then performed to get $PS(x(t))$ which is also also the signal for the first iteration. Iteration is then performed by adding to the

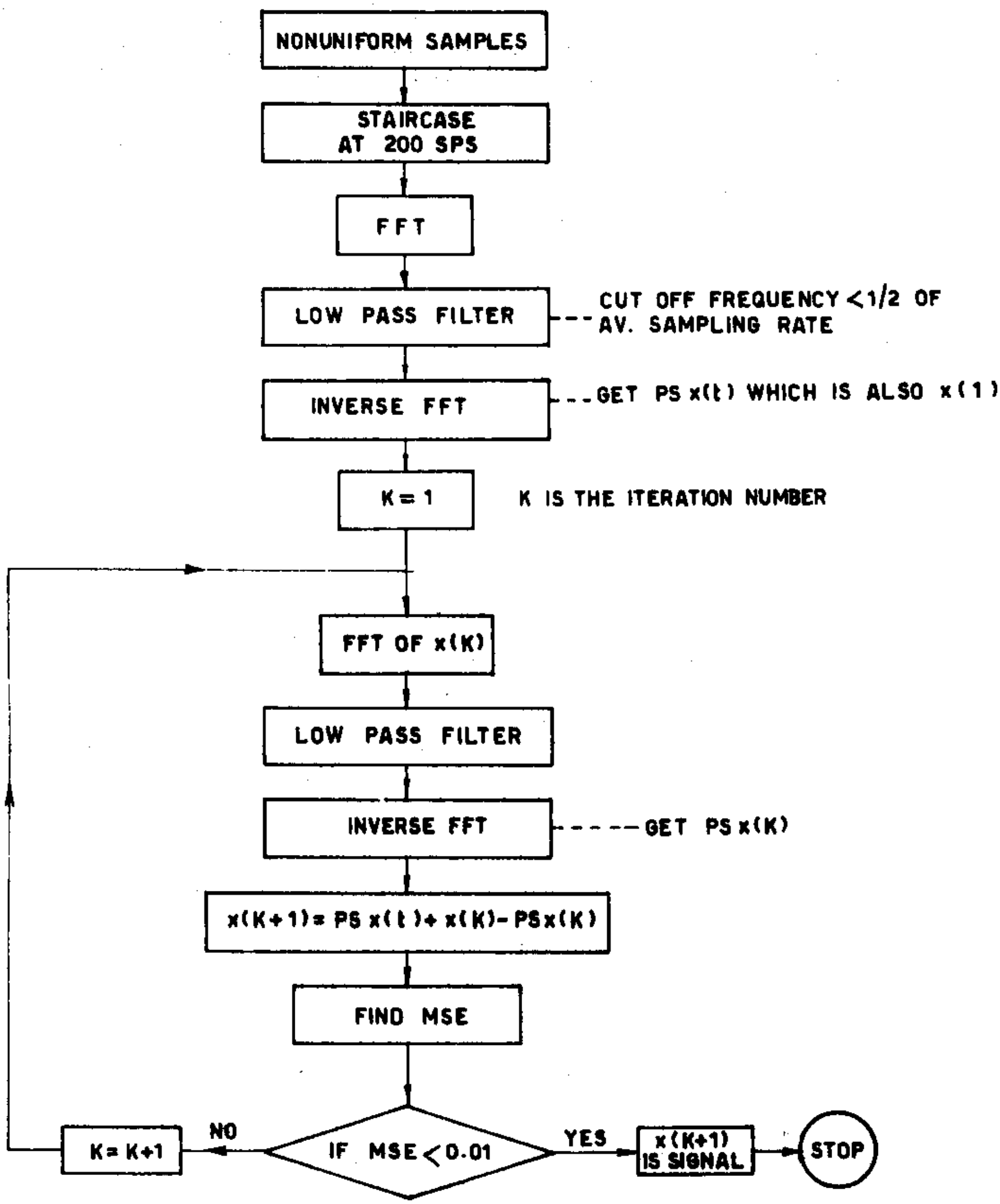

FIG. 1_ FLOWCHART FOR RECOVERY OF BAND LIMITED SIGNAL FROM NONUNIFORM SAMPLES FOR $\lambda = 1$

$PS(x(t))$,the difference between the recovered signal of earlier iteration and its band operated signal. After each iteration mean square error (MSE) between x_{k+1} and x_k is determined. MSE is defined as :

$$MSE = \frac{\sum_i (\epsilon^i - \bar{\epsilon})^2}{ndata - 1} \tag{8}$$

where i is sample points and

$$\epsilon = x_{k+1} - x_k \tag{9}$$

and

$$\bar{\epsilon} = \frac{\sum_i \epsilon^i}{ndata} \tag{10}$$

The desired signal is assumed to be achieved and the iteration is assumed to have converged when value of MSE becomes less than 0.01. The computer program developed for this alogorithm is interactive in nature and is quite flexible with regard to the choice of various characteristics like cutoff frequency, number of poles of filter, maximum number of iterations to be used etc.

RECOVERY OF SINE WAVES & EARTHQUAKE SIGNALS

This computer program was first used to recover a sinewave having frequency of 10 *Hz*. The sine wave was defined by nonuniform samples at an average sampling rate of 50 *SPS*. The nonuniform samples were distributed so as to satisfy Inequality 2. The distribution of samples within the range of Inequality 2 was done randomly by generating random numbers. This software was then used to recover the 10 *Hz* signal at 200 *SPS* and with a cut off frequency of 25 *Hz*. The mean square error (MSE) of the recovered signal was determined after each iteration. After 10 iterations MSE less than 0.01 was obtained. The Fourier spectra of the signal so obtained after 10 iterations is shown in Figure 2. A linear interpolation was done on the same nonuniform samples of 10 *Hz* to obtain the sequence at 200 *SPS*. The Fourier spectra of sequence obtained after linear interpolation is shown in Figure 3. The comparison of Figures 2 and 3 clearly indicates that the linear interpolation introduces high frequency jitters and reduces energy from the required frequency. Whereas, the band limited interpolation has almost zero noise in frequencies greater than 25 *Hz* (stop band) and has smaller noise in the pass band.

The computer program was then used to recover sine wave of 10 *Hz* which was defined by nonuniform samples at an average sampling rate of 50 *SPS*. However, now nonuniform samples were distributed randomly without enforcing the restriction of Inequality 2. In this case MSE less than 0.01 was obtained after 25 iterations. The Fourier spectra of signal of 10 *Hz* at 200 *SPS* so obtained after 25 iterations is shown in Figure 4. A linear interpolation was also done on the random nonuniform samples to obtain the sequence at 200 *SPS* and Figure 5 shows its Fourier spectra. Comparison of Figures 4 and 5 indicates results as in the case of stable sampling set except that random nonuniform samples need more iterations to recover the signal.

Next in the series, an attempt was made to recover 30 *Hz* sinewave which was defined by nonuniform samples with an average sampling rate of 50 *SPS* (less than the Nyquist samplig rate). The software could not recover the sine wave because the iteration process did not converge.

Next this computer program was used to recover the signal from uncorrected earthquake accelerogram recorded from an analog accelerograph and digitized at nonuniform interval on a semi-automatic digitizer. For this purpose an accelerogram obtained from Uttarkashi earthquake

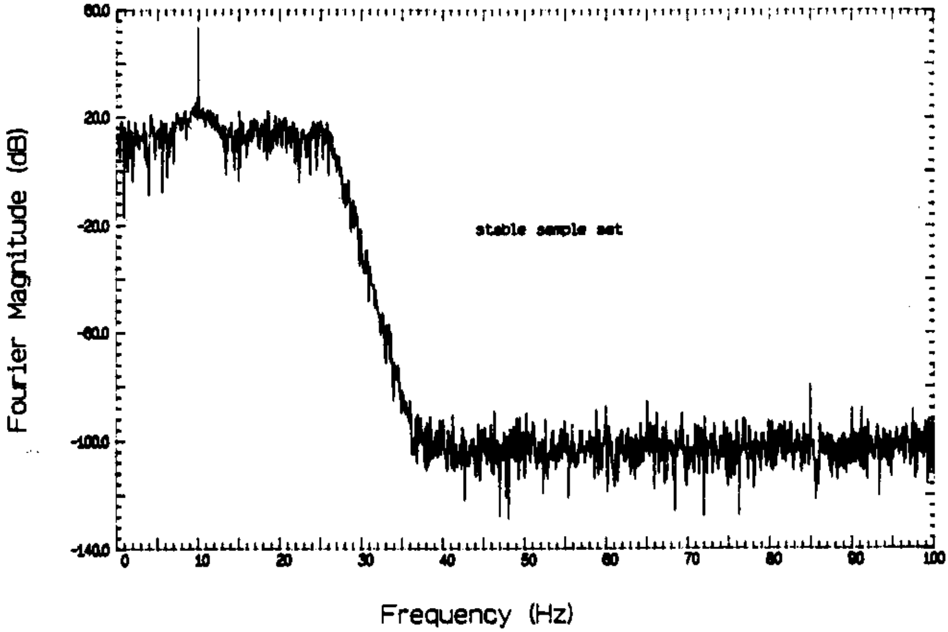

Figure 2: Band limited sine wave from nonuniform samples.

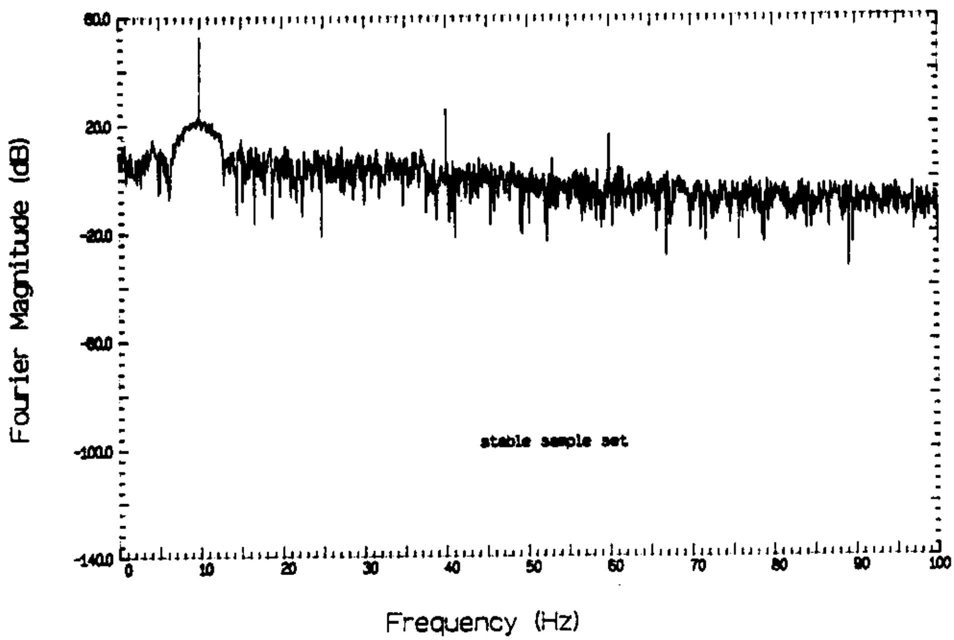

Figure 3: Linearly interpolated sine wave from nonuniform samples.

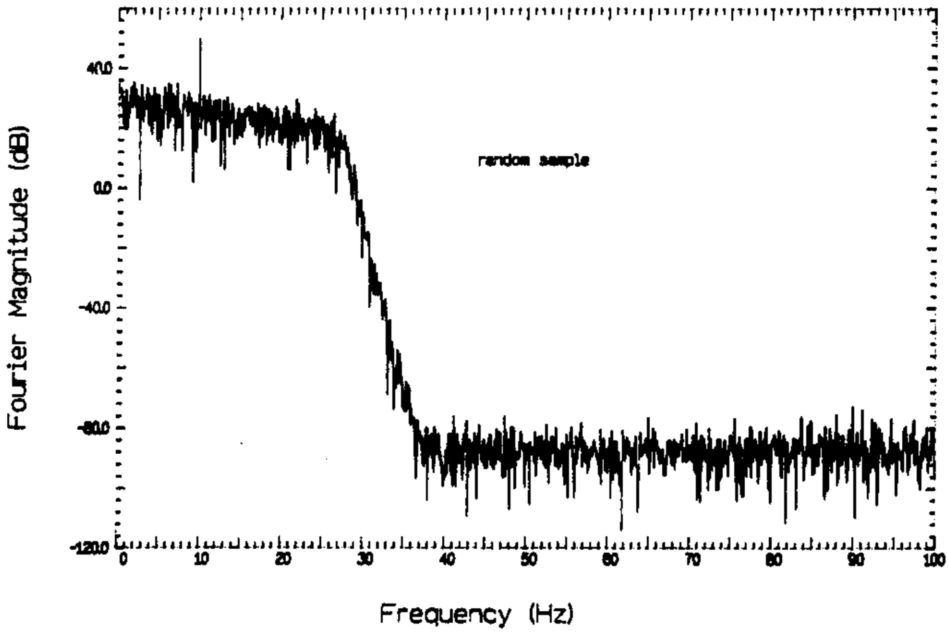

Figure 4: Band limited sine wave from nonuniform samples.

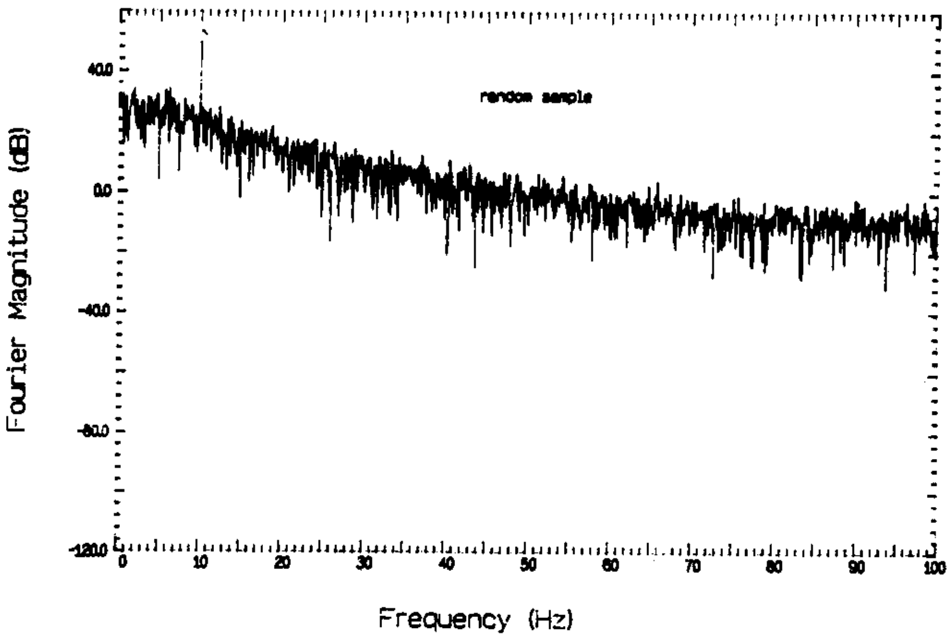

Figure 5: Linearly interpolated sine wave from nonuniform samples.

of October 20, 1991 was used. This accelerogram had a duration of 37.13 seconds and had 2902 sample points. Thus, the average sampling rate of the accelerogram is much more than the Nyquist rate corresponding to 25 *Hz*. This accelerogram was recovered at 200 *SPS* through the proposed iteration scheme after 25 iterations for MSE less than 0.01. Fourier spectra of the recovered accelerogram is shown in Figure 6. The uncorrected accelerogram at nonuniform samples was linearly interpolated to get the sequence at 200 *SPS* and Figure 7 shows Fourier spectra of linearly interpolated sequence. Comparison of Figures 6 and 7 show that the undesired high frequency jitters exist in linearly interpolated sequence and even in the frequency band upto 25 *Hz* there is difference in the two spectra which is due to distortion of signal during linear interpolation due to introduction of higher frequencies and effect of aliasing.

REVIEW OF DATA OF SOME PAST EARTHQUAKES

Several published uncorrected accelerograms were reviewed in the process of above study. It is found that most of uncorrected accelerograms of earthquakes which occured before 1970 were sampled at an average rate of less than 50 *SPS* (Nyquist rate corresponding to signal of 25 *Hz*) and several uncorrected accelerograms of earthquakes which occured after 1970 also have this problem. In such a situation any attempt made to obtain information about frquencies more than half of the average sampling rate will only distort the entire band of frequencies due to effect of aliasing. This aspect has been apparently ignored so far and uncorrected accelerograms were linearly interpolated and a cutoff frequency of 25 *Hz* was used in the processing of these accelerograms. In other words the processing methodology adopted till now had attempted to obtain information in the frequency range which had been preluded by the digitization process. Such an attempt corrupts the entire frequency range. In this part of the work, records of El Centro earthquake of May 18, 1940 and Parkfield earthquake of June 27, 1966 are examined and band limited signal was recovered from these uncorrected accelerograms. This band limited uncorrected accelerogram was processed for instrument correction and was band-pass filtered with a lower cut-off frequency of 0.07 *Hz* and upper cut-off frequency of little less than half of average sampling rate of uncorrected data. Response spectra of this time history was calculated and compared with response spectra of time history obtained with conventional method of linear interpolation with cut-off frequency of 25 *Hz*.

El Centro Accelerogram

The N-S component of the uncorrected accelerogram of Imperial valley earthquake of May 18, 1940 recorded at El Centro (EERL 70-20, File 1) is examined here. This accelerogram has 985 digitized points and the duration of the record is 53.732 *sec*. This amounts to an average sampling rate of 18.33 *SPS* and therefore largest frequency which can be recovered from such data is 9 *Hz*. Band limited interpolation as discussed before was performed to recover the signal at 200 *SPS* with cut-off frequency of low-pass filter taken as 9 *Hz*. The natural frequency of accelerometer was 10 *Hz* and damping was 55.2% of critical. Instrument correction was then performed with these parameters and signal was band-pass filtered with a lower cut-off frequency of 0.07 *Hz* and upper cut-off frequency of 9 *Hz*. The corrected accelerogram so obtained is the best description of ground acceleration which can be obtained with the available uncorrected accelerogram. The response spectra (for 2% damping) using algorithm of Nigam and Jennings (1969) calulated for this time history alongwith response spectra calculated for time history obtained through conventional method is given in Figure 8. Figure 9 gives the same response spectra with natural frequency in abscissa. The comparison in Figures 8 and 9 show that the

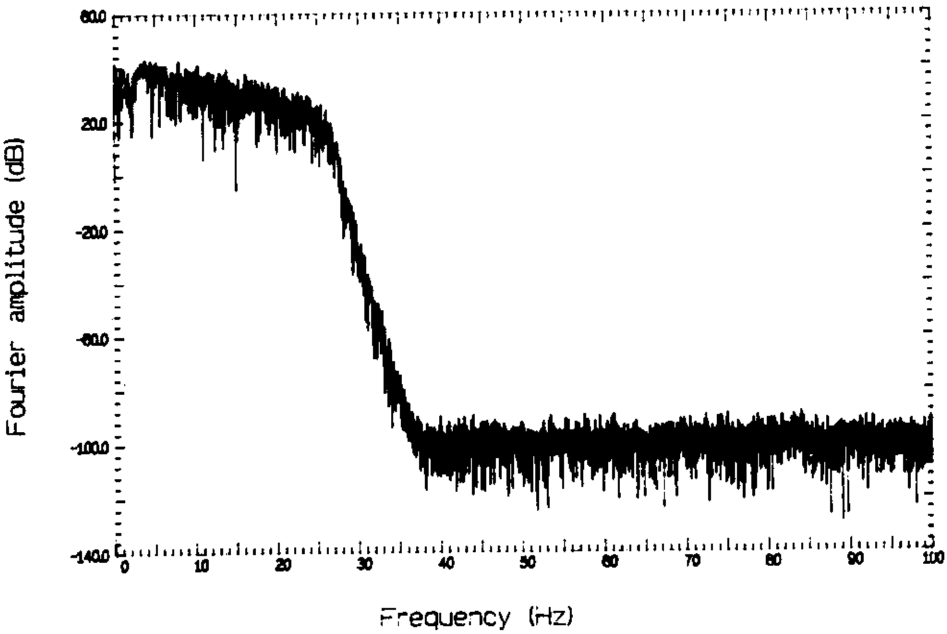

Figure 6: Band limited uncorrected Uttarkashi accelerogram.

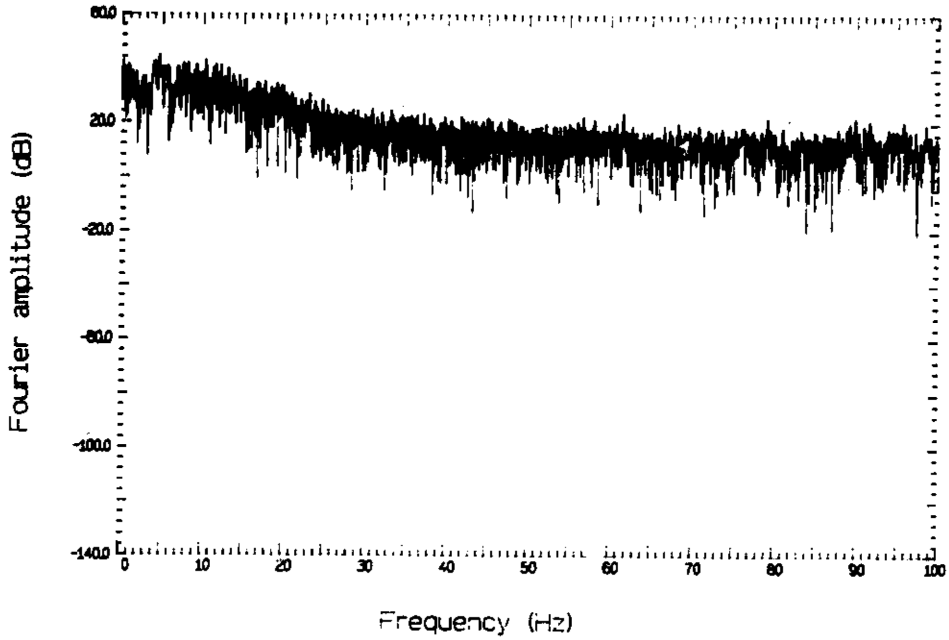

Figure 7: Linearly interpolated uncorrected Uttarkashi accelerogram.

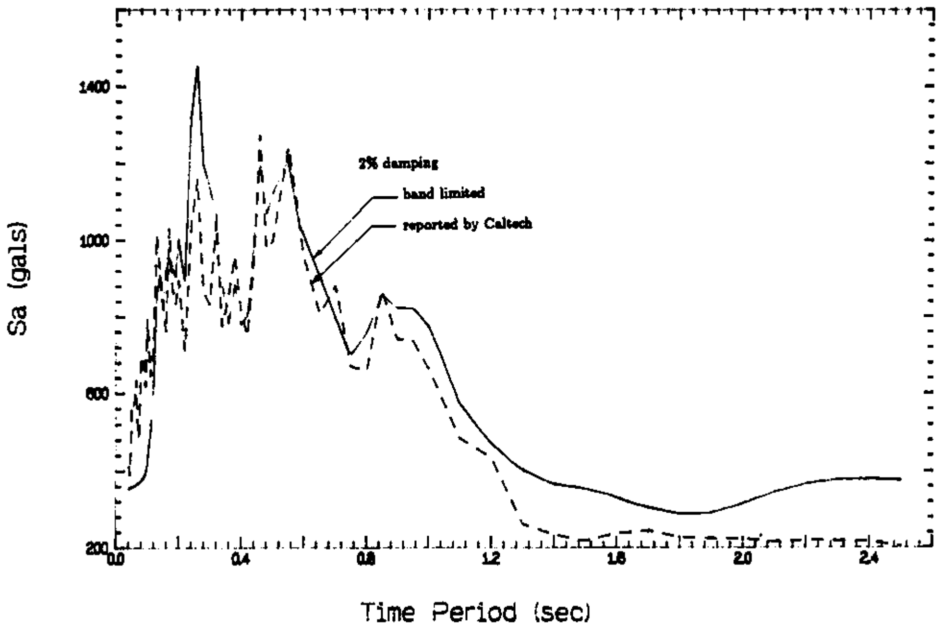

Figure 8: El Centro (NS) response spectra – conventional (obtained from Caltech report Vol II A, record no. 001) and bandlimited.

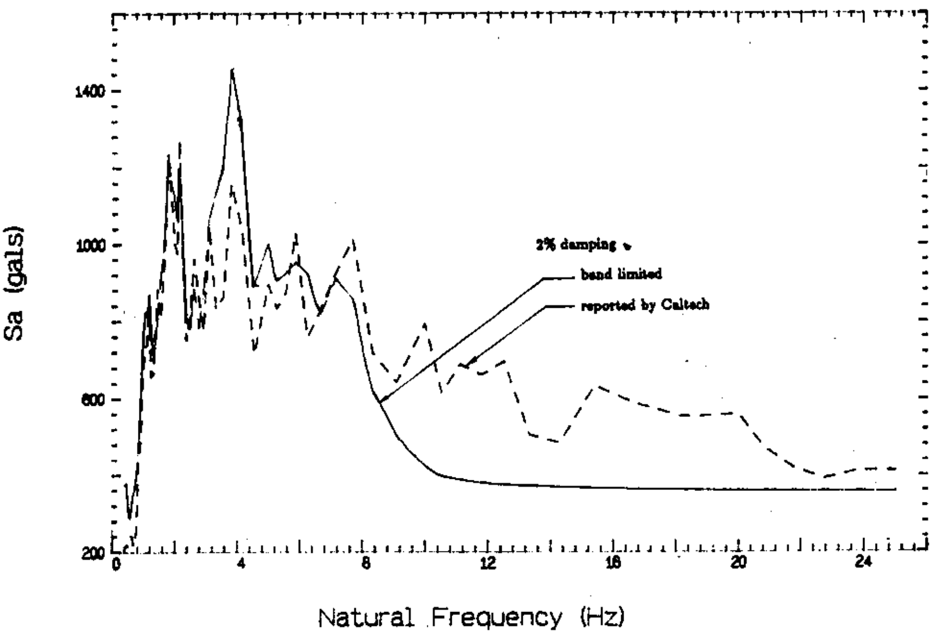

Figure 9: El Centro (NS) response spectra – conventional (obtained from Caltech report Vol II A, record no. 001) and bandlimited.

conventional response spectra of El Centro N-S component has been underestimated in periods more than 0.25 seconds. However, at frequencies more than 9 *Hz* response spectra obtained through conventional method overestimates the spectral values. In fact Figure 9 clearly shows the effect of aliasing in the conventional time history due to which overestimation of spectral values in the high frequency range has occurred.

Parkfield Accelerogram

Uncorrected accelerogram of Parkfield earthquake of June 27, 1966 recorded at Cholame, Shandon (EERL 70-21, File 40) is studied next. The N05W component of this accelerogram has 1010 digitized points and the duration of the record is 44.01 *sec*. This amounts to an average sampling rate of 22.95 *SPS* and therefore largest frequency which can be recovered from such data is 11 *Hz*. Band limited interpolation as discussed before was performed to recover the signal at 200 *SPS* with cut-off frequency of low-pass filter taken as 11 *Hz*. The natural frequency of accelerometer was 19.61 *Hz* and damping was 37% of critical. The signal was processed as mentioned in the earlier subsection. The response spectra (for 2% damping) calulated for this time history alongwith response spectra (for 2% damping) calculated for time history obtained through conventional method is given in Figure 10. Figure 11 gives the same response spectra with frequency in abscissa. This time the results are found to be almost similar with a little overestimation of conventional method in high frequency range. The conclusions which can be drawn from this study is that in certain set of data the effect of aliasing as well as cut-off frequency may not get reflected in the response spectra as in this case. But in the other set of data this effect can substantially change the characteristics of the response spectra as in the case of El Centro accelerogram. However, the point remains that the basic principle of digital signal processing is flouted by attempting to extract information about the frequency contents which are larger than half of average sampling rate and also in some cases such attempts can make good amount of difference in the final results.

CONCLUSIONS

1. Based on the Nyquist theorem and definition of stable sampling set, guidelines can be evolved to operators performing digitization of accelerograms on semi automatic digitizers as discussed in section on stable sampling set.

2. Some of the published uncorrected accelerograms were reviewed and it is found that several of them were digitized at sampling rate lower than required for recovering frequencies upto 25 *Hz*. In addition to El Centro and Parkfield accelerograms discussed, many other past accelerograms have this problem. For example, in EERL Report no. 73-28 (1973) record number 70.009 having duration of about 42 seconds has 1513 sample points for North-South component (average sampling rate $\simeq$ 36 *SPS*), 1594 sample points for East-West component (average sampling rate $\simeq$ 38 *SPS*) and 1805 sample points for Vertical component (average sampling rate $\simeq$ 43 *SPS*). The largest frequencies which can be recovered from such data is 18 *Hz* for North-South component, 19 *Hz* for East-West component and 21.5 *Hz* for vertical component. There are several other similar examples in the published reports of uncorrected accelerograms of this as well as other earthquakes. The point that emerges is that because of prevalent practice of performing linear interpolation, the operators who digitize the records, consider it appropriate to pick only peaks and troughs of the record with few additional points in between which is likely to yield limited information only.

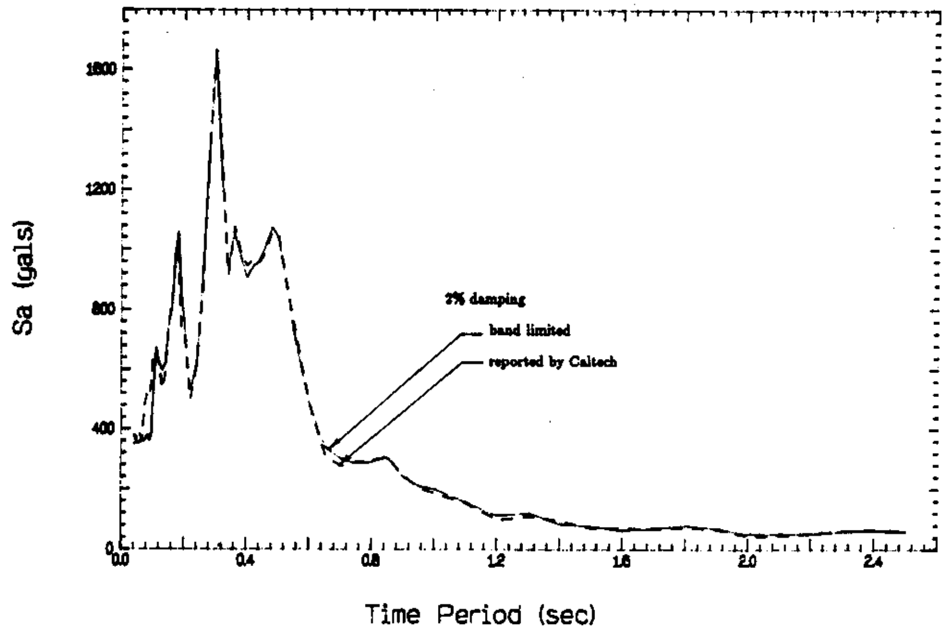

Figure 10: Parkfield (N05W) response spectra – conventional (obtained from Caltech report Vol II B, record no. 034) and bandlimited.

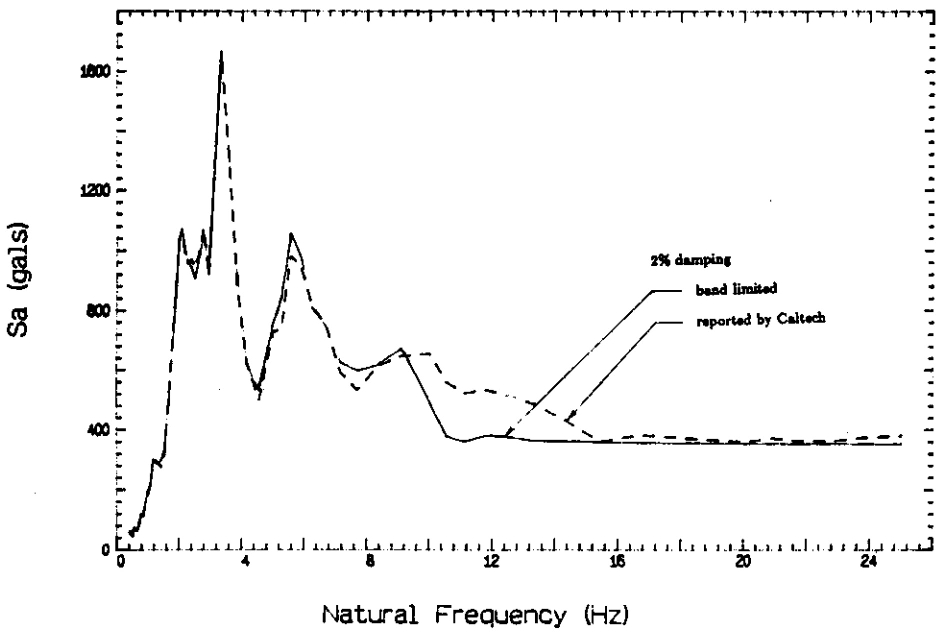

Figure 11: Parkfield (N05W) response spectra – conventional (obtained from Caltech report Vol II B, record no. 034) and bandlimited.

3. Examples on recovery of sine waves clearly demonstrate that linear interpolation introduce noise in the high frequency range at the cost of reducing the strength of actual signal whereas the proposed band limited interpolation preserves the frequency content of the data. These features remain unchanged whether one uses stable or random sample sets as long as Nyquist criterion of average sampling rate is satisfied. However, when stable sampling set is used the convergence speed of the iteration process increases.

4. From the above conclusions it can be deduced that the paper/film speed of accelerograph becomes important in semi-automatic digitizers. A higher paper speed will make it easier for the opertor to perform digitization with greater average sampling rate. For example the paper speed of 1 *cm/sec* available in SMA 1 accelerograph manufactured by Kinemetrics, USA does not appear to be sufficient for digitising an average of 50 points per centimeter. In this regard RESA V of University of Roorkee is better placed as it has a paper speed of 2 *cm/sec*.

5. El Centro and Parkfield accelerograms examined, show that conventional processing corrupted the corrected accelerogram of El Centro earthquake but the same is not reflected in the Parkfield accelerogram. This could be due to concentration of energy in high or low frequency range or due to aliasing which may have neutralised the effect of linear interpolation. The authors are investigating this further.

6. It appears that the best solution with regard to the use of analog accelerograms is to employ automatic scanners which automatically perform digitisation at constant time interval at sufficiently large sampling rate. However, the fact remains that large number of past accelerograms have been digitized on semi-automatic digitizers and information contained in it is valuable input for earthquake engineering. In some places semi-automatic digitizers are still commonly used hence an accurate method of processing such data is needed.

REFERENCES

1. **Lam H.Y.F. (1979)** —Analog and Digital Filters: Design and Realization. *Prentice-Hall, Inc., Englewood, New Jersey 07632*

2. **Marvasti, F. (1987)** —A unified approach to zero crossings and nonuniform sampling of single and multidimensional siganals and systems. *Oak Park, IL: Nonuniform, first ed.*

3. **Marvasti, F., Analoui, M.** and **Gamshadzahi, M. (1991)** —Recovery of Signals fro Non Uniform Samples using Iterative Methods. *IEEE Transaction on Signal Processing,* **Vol. 39, No. 4.**

4. **Nigam, N.C.** and **Jennings, P.C. (1969)** —Calculation of Response Spectra from Strong Motion Earthquake Records. *Bulletin of the Seismological Society of America,* **Vol. 59, No. 2**, pp909-922

5. —(1970)Strong Motion Earthquake Accelerograms; Vol.I Uncorrected accelerograms. *Report of Earthquake Engineering Research Laboratory, California Institute of Technology,* **Report no. EERL 70-20, Vol. I, Part A**

6. —(1970)Strong Motion Earthquake Accelerograms; Vol.I Uncorrected accelerograms. *Report of Earthquake Engineering Research Laboratory, California Institute of Technology,* **Report no. EERL 70-21, Vol. I, Part B**

7. —(1973)Strong Motion Earthquake Accelerograms; Vol.I Uncorrected accelerograms. *Report of Earthquake Engineering Research Laboratory, California Institute of Technology,* **Report no. EERL 73-28, Vol. I, Part W**