ICSV14
Cairns • Australia
9-12 July, 2007
14th International Congress on Sound & Vibration

FRACTIONAL FOURIER TRANSFORMS FOR SPECTRAL ANALYSIS OF MUSIC SIGNALS

Charu[1 *], V.G.Das[2 †], V.Prem Pyara[3 ‡], V.Prem Kumari[4 §]

[1]Jaypee Institute of Engineering and Technology, Raghogarh, India

[2]Dayalbagh Educational Institute, Agra, India

[3]Department of Electrical Engineering, Dayalbagh Educational Institute, Agra, India

[4]Department of Music, Dayalbagh Educational Institute, Agra, India

[*]pathak_charu@rediffmail.com
[†]admin@dei.ac.in
[‡]prempyara@hotmail.com
[§]dbei@vsnl.com

Abstract

A new technique for spectral analysis is proposed for application to music signals. The frequency transform presented in this work eliminates the need for multiplications, by requiring the sole use of additions making the ASIC (Application Specific Integrated Circuit) implementation simpler. The transform is generalized and can be applied to any band limited spectral analysis. The proposed Fractional Fourier Transform shows almost no error in the analysis when compared to standard Discrete Fourier Transform.

1. INTRODUCTION

The most efficient technique often used for spectral analysis is Fast Fourier Transform. The N point FFT requires $Nlog_2N$ multiplications and additions of complex numbers. In order to avoid multiplications, it is suggested that square waves may be used in place of sinusoidal waves for spectral estimation. The advantage of using square waves in terms of computation has been exploited earlier as Haar Transforms or Walsh Transforms etc [7, 8, 1]. In this work, the aim is to obtain exact Fourier spectrum for a specified frequency range using square waves in place of sinusoidal functions. The square waves consist of amplitudes of +1 and -1 only and thus avoid any multiplications required in FFT. Poorman's Transform developed by Michael. P. Lamoureux also eliminates multiplications [2, 3, 4]. If the signal is integrated initially, the computations get reduced as the multiplication and integration of half cycle of square wave reduces to simple subtraction of the values at the edges of the square wave [10]. A composite square wave is proposed in this work, so as to approximate the sine wave. This composite square wave is realized by the summation of the fundamental square wave and its harmonics. It is shown that such a wave can extract exact spectral amplitudes within a certain range of the spectrum.The technique is suitable for pitch analysis of music signals as the analysis is limited to an octave bandwidth. It is applied to Indian Classical Music for note detection.

2. ESTIMATION OF MINIMUM QUALITY FACTOR FOR PITCH PERCEPTION IN INDIAN MUSIC

The ears of experienced musicians identify small deviations in tune quite easily. The scale of Indian Music divides an octave into seven swaras and further divides them into srutis. According to Dr.S.S.Bhave, a scientist cum musician working with CSC group at TIFR, Mumbai there are three intervals that are represented by a sruti in Indian Music system [5]. The intervals are $\frac{81}{80}$, $\frac{25}{26}$, $\frac{256}{243}$. The smallest of the sruti ratios so identified is $\frac{81}{80}$. This is similar, as resolving $1.25Hz$ in $100Hz$. The resolution power of good musician is considered to be equal to 1Hz at 100Hz [6]. This is equivalent to a per unit resolution of 0.01. The Fractional Fourier Transform (FrFT) technique developed here is optimized for this resolution and uses this fact for reducing computations.

3. COMPOSITE SQUARE WAVE APPROXIMATION OF A SINE WAVE

According to Fourier series estimation, square wave can be expressed as a sum of sine and cosine functions. Thus, a sine wave may be expressed as a sum of square waves, consisting of a fundamental frequency component and its frequency multiples. The first order approximation of sine wave is a symmetrical square wave oscillating between the values $+1$ and -1, where the $+1$ value corresponds to positive half cycle of $\sin(\omega_0 t)$ and -1 corresponds to its negative half cycle. Mathematically, the set of harmonics of these square wave functions may be defined using signum functions as,

$$\mathrm{S_e}(n\omega_0 t) = \mathrm{sgn}[\cos(n\omega_0 t)] \tag{1}$$

and

$$\mathrm{S_o}(n\omega_0 t) = \mathrm{sgn}[\sin(n\omega_0 t)]. \tag{2}$$

The set of sine and cosine functions form an orthogonal set and any function in time domain can be expressed as a sum of orthogonal projections. Unfortunately, the square wave harmonics are not orthogonal. The accuracy of approximation improves by increasing the number of harmonics of the square wave that are taken into consideration. Here, the basis for this approximation is worked out. The Fourier expansion of the square wave $\mathrm{S_e}(\omega_0 t)$ is given by

$$\begin{aligned} s(t) &= \tfrac{4A}{\pi}\{\cos(2\pi f_0 t) - \tfrac{1}{3}\cos(3.2\pi f_0 t) + \ldots + \tfrac{(-1)^n}{2n+1}\cos((2n+1)2\pi f_0 t) + \ldots\} \\ &= \tfrac{2A}{\pi}\sum_{n=-\infty}^{n=+\infty} \tfrac{(-1)^n}{(2n+1)} e^{\{j(2n+1)\omega_0 t\}}. \end{aligned} \tag{3}$$

Thus, $\cos(\omega_o t)$ may be approximated as $\frac{\pi}{4}\mathrm{S_e}(\omega_o t)$. The error term will consist of third, fifth and higher odd harmonics of $\omega_o t$. This may be written mathematically as,

$$\begin{aligned} \cos(\omega_o t) = \; & \tfrac{\pi}{4}\mathrm{S_e}(\omega_o t) + \tfrac{1}{3}\cos(3\omega_o t) - \tfrac{1}{5}\cos(5\omega_o t) + \ldots \\ & + \tfrac{(-1)^{n+1}}{2n+1}\cos((2n+1)\omega_o t) + \ldots . \end{aligned} \tag{4}$$

Approximating $\cos(\omega_0 t)$ by $\frac{\pi}{4}\mathrm{S_e}(\omega_0 t)$ results in an error $e(t)$. This error is given by,

$$e(t) = \frac{1}{3}\cos(3\omega_o t) - \frac{1}{5}\cos(5\omega_o t) + \ldots + \frac{(-1)^{n+1}}{2n+1}\cos((2n+1)\omega_o t) + \ldots . \tag{5}$$

The mean square value or the power content of $\cos(\omega_o t)$ is $\frac{1}{2}$ while that of $\frac{\pi}{4}\mathrm{S_e}(\omega_o t)$ is $\frac{\pi^2}{16}$. Thus the power content of error term is 0.114 $\{i.e.,\ (\frac{\pi^2}{16} - \frac{1}{2})\}$, which is approximately 22.8%. The power of the error function is 0.114. But this power subtracts from that of the square wave $\frac{\pi}{4}\mathrm{S_e}(\omega_o t)$. This is because $\mathrm{S_e}(\omega_o t)$ is not orthogonal to the terms in the error function. The error is reduced by replacing $\cos(3\omega_o t)$ by an equivalent square wave and higher harmonics of $\cos(3\omega_o t)$. This expression is obtained by replacing ω_o by $3\omega_o$ in the eqn. 4, as

$$\begin{aligned}\cos(3\omega_o t) = \ & \tfrac{\pi}{4}\mathrm{S_e}(3\omega_o t) + \tfrac{1}{3}\cos(9\omega_o t) - \tfrac{1}{5}\cos(15\omega_o t) + \ldots \\ & + \tfrac{-1^{n+1}}{2n+1}\cos((2n+1)\omega_o t) + \ldots .\end{aligned} \tag{6}$$

Combining eqns. 4 and 6, we get

$$\begin{aligned}\cos(\omega_o t) \ & = \tfrac{\pi}{4}\{\mathrm{S_e}(\omega_o t) + \tfrac{1}{3}\mathrm{S_e}(3\omega_o t)\} - \tfrac{1}{5}\cos(5\omega_o t) \\ & + \tfrac{1}{7}\cos(7\omega_o t) + \tfrac{1}{11}\cos(11\omega_o t) - \tfrac{1}{13}\cos(13\omega_o t) \\ & - \tfrac{1}{17}\cos(17\omega_o t) + \tfrac{1}{19}\cos(19\omega_o t) + \tfrac{1}{23}\cos(23\omega_o t) - \ldots\}.\end{aligned} \tag{7}$$

Note that the multiples of $3\omega_o t$ get exactly canceled out. The power of the signal $\cos(\omega_o t)$ is 0.5 while that of square wave function is $\frac{\pi^2}{18}$. Thus the power in the error term gets reduced to nearly 0.046. This approximation may be extended by replacing $\cos(5\omega_o t)$ and $\cos(7\omega_o t)$ and other higher harmonics also by the square wave functions $\mathrm{S_e}(5\omega_o t)$ and $\mathrm{S_e}(7\omega_o t)$ and corresponding harmonic components leading to a more exact expression as

$$\cos(\omega_0 t) = \frac{\pi}{4}\{\mathrm{S_e}(\omega_o t) + \frac{1}{3}\mathrm{S_e}(3\omega_o t) - \frac{1}{5}\mathrm{S_e}(5\omega_o t) + \frac{1}{7}\mathrm{S_e}(7\omega_o t)\} + e(t) \tag{8}$$

. The power of the function, the power of the approximated square wave and that of the error term are $0.5,\ 0.520335$ and 0.020335 respectively. The power of the error term is thus reduced. It can be reduced further by replacing the higher harmonic terms with the corresponding square wave components.Since the signal is band-limited, it is not necessary to evaluate all the terms in the expansion. The following section gives the expressions for evaluating exact Fourier spectral amplitudes using the composite square waves in place sinusoidal function.

4. ESTIMATION OF SQUARE WAVE FOURIER COEFFICIENTS

To evaluate square wave Fourier coefficients the term $\cos(n\omega_o t)$ in Fourier coefficients $\{a_n\}$ is replaced by first order square wave approximation $\frac{\pi}{4}\mathrm{S_e}(n\omega_o t)$. The resulting coefficients are given by

$$a_0'' = a_0 \tag{9}$$

,

$$a_n'' = \frac{2}{T}\int_0^T x(t)\{\frac{\pi}{4}\mathrm{S_e}(n\omega_o t)\}dt \,. \tag{10}$$

and

$$b''_n = \frac{2}{T}\int_0^T x(t)\{\frac{\pi}{4}\mathrm{S_o}(n\omega_o t)\}dt\ . \tag{11}$$

To find a''_n in terms of $\{a_n\}$, the square wave $\mathrm{S_e}(n\omega_o t)$ in eqn. 10 is replaced by its Fourier components as

$$\begin{aligned} a''_n = & \ \tfrac{\pi}{4}\tfrac{2}{T}\textstyle\int_0^T x(t)\{\cos(n\omega_o t) - \tfrac{1}{3}\cos(3n\omega_o t) + \ldots \\ & +(-1)^m \tfrac{1}{(2m+1)}\cos((2m+1)\omega_o t) + \ldots\}dt \\ = & \ \ a_n - \tfrac{a_{3n}}{3} + \tfrac{a_{5n}}{5} - \ldots + \tfrac{(-1)^m a_{(2m+1)n}}{(2m+1)} + \ldots . \end{aligned} \tag{12}$$

The result shows that a''_n differs from a_n by a fraction of the amplitudes of third, fifth and higher odd harmonics of $n\omega_o$ present in $x(t)$. This approximation has infinite terms. In practice one has to limit the sequence upto certain number of terms. The truncated sequences are denoted by $\cos'(n\omega_0 t)$ and $\sin'(n\omega_0 t)$. As an example, the terms upto 7^{th} harmonic only are considered in the truncated sequences. The expression for $\cos'(n\omega_0 t)$ with terms upto seventh harmonic come out as,

$$\begin{aligned} \cos'(n\omega_0 t) = & \ \tfrac{\pi}{4}\{\mathrm{S_e}(n\omega_o t) + \tfrac{1}{3}\mathrm{S_e}(3n\omega_o t) \\ & -\tfrac{1}{5}\mathrm{S_e}(5n\omega_o t) + \tfrac{1}{7}S_e(7n\omega_o t)\} \end{aligned} \tag{13}$$

The error terms contain 11^{th} and higher harmonics only. All the lower harmonics either get canceled or are taken into account by the square wave terms. The exact expressions of $\cos'(n\omega_0 t)$ come out as,

$$\begin{aligned} \cos'(n\omega_0 t) = & \ \cos(n\omega_o t) - \tfrac{1}{11}\cos(11n\omega_o t) \\ & +\tfrac{1}{13}\cos(13n\omega_o t) + \tfrac{1}{17}\cos(17n\omega_o t) + \ldots \end{aligned} \tag{14}$$

When square wave expression of $\cos'(n\omega_0 t)$ given in eqn. 13 is used as an approximation of $\cos(n\omega_o t)$ the resulting coefficients a'_n are given by

$$\begin{aligned} a'_n = & \ \tfrac{2}{T}\textstyle\int_0^T x(t)\left[\tfrac{\pi}{4}\left\{S_e(n\omega_0 t) + \tfrac{1}{3}S_e(3n\omega_0 t)\right.\right. \\ & \left.\left. -\tfrac{1}{5}S_e(5n\omega_0 t) + \tfrac{1}{7}S_e(7n\omega_0 t)\right\}\right] dt \end{aligned} \tag{15}$$

These coefficients differ from their true values a_n due to the harmonic terms neglected. Their relation is derived from the relation given in eqn. 14 as

$$a'_n = a_n - \frac{1}{11}a_{11n} + \frac{1}{13}a_{13n} + \frac{1}{17}a_{17n} + \ldots \tag{16}$$

. Similarly,

$$b'_n = b_n - \frac{1}{11}b_{11n} - \frac{1}{13}b_{13n} - \frac{1}{17}a_{17n} + \ldots\ . \tag{17}$$

The eqns. 16 and 17 show that the estimated value, a'_n is corrupted by the addition of eleventh, thirteenth and higher harmonics. For a band-limited voice signal, if the highest frequency component is less than $11n\omega_o$, then all the components $a_{11n}, a_{13n}, a_{17n}, \ldots$ etc. will be zero and the calculated value a'_n will be exactly equal to a_n. In case the highest frequency of $x(t)$ has harmonic components greater than $11n\omega_o$, more higher harmonic terms may be added in the square wave expansion. If the composite square wave signal includes $(2m-1)^{th}$ square wave harmonic, the error term includes $(2m+1)^{th}$ harmonics and other higher harmonics. But increasing the square wave harmonic terms beyond the highest frequency of signal does not increase the accuracy in any way. The number of square wave harmonic terms to be considered so that one can obtain error free spectral analysis for the pitch estimation of a vocal music signal depends on the spectral range of analysis. This is discussed in the following section.

5. NUMBER OF HARMONICS IN COMPOSITE SQUARE WAVES FOR THE ANALYSIS OF A VOCAL MUSIC SIGNAL

When the composite square waves are used in place of sinusoidal functions, the exact spectrum can be obtained in the limited range of frequencies only as the case of vocal music signal.The number of square wave terms to be included in the realization of the composite square wave depends upon this range. Let this range be from f_1Hz to f_2Hz. For audio signal analysis, this range starts from a frequency of $100Hz$ and extends upto a range of few thousands of Hz for a resolution of 1%, as mentioned above. For the estimation of the minimum number of square wave harmonics required for pitch analysis of such a signal, consider a bandlimited signal with a maximum frequency of f_hHz. When the composite square wave having square wave terms upto $(2m-1)^{th}$ harmonic is used to estimate the spectrum, several image spectra are produced which corrupt the result. These image spectra are produced because the $(2m+1)^{th}$ and higher harmonic terms are not compensated in the composite square wave as its takes care of $(2m-1)^{th}$ harmonic only. The frequency range of these image spectra extend from d.c. to $\frac{f_h}{2m+1}$. The coefficients corresponding to the region $\frac{f_h}{2m+1}$ to f_h is not corrupted by image spectra. The desired range of spectral estimation which extends from f_1Hz to f_2Hz should lie in this uncorrupted region. The value of m in the composite square wave be such that $\frac{f_h}{2m+1} < f_1$. In order to construct the simplest composite square wave, one has to minimize the value of m. This leads to the condition,

$$2m-1 < \frac{f_h}{f_1} < 2m+1. \tag{18}$$

That is, for a given ratio $\frac{f_h}{f_1}$, the order of square wave harmonic terms to be included in the composite square wave should be $2m-1$, which is constrained by the eqn. 18 given above. The above analysis is given for the continuous time signals. In practice, this has to be implemented on a computer in sampled domain. The necessary algorithm is given below.

6. COMPOSITE SQUARE WAVE ESTIMATION FOR ERROR FREE FOURIER COEFFICIENTS

The signal is first sampled at a certain fixed rate of f_s samples per second and the values are stored in an array. Due to non-stationary nature of signal, we subdivide the array into packets of suitable length. This process is similar to that adopted in the evaluation of Short Time Fourier Transform. Let the length of this packet be N. The integrals similar to those given in eqn. 15 are to be evaluated using the elements of this array. The integrals are converted into summation.Instead of evaluating the entire expression consisting of all harmonic terms, it is suggested that coefficients should be first evaluated using a single square wave and compute the composite square wave coefficients later. The single square wave coefficients are termed as a''_n and b''_n

(Refer eqn. 10). The integral expressions are,

$$a''_n = \frac{2}{T}\int_0^T x(t)\{\frac{\pi}{4}S_e(n\omega_0 t)\}dt \tag{19}$$

and

$$b''_n = \frac{2}{T}\int_0^T x(t)\{\frac{\pi}{4}S_o(n\omega_0 t)\}dt. \tag{20}$$

The above expressions can be modified for the analysis in sampled domain as,

$$a''_n = \frac{2}{N}\frac{\pi}{4}\sum_{k=0}^{N-1} x(k)S_e(\frac{2\pi nk}{N}) \tag{21}$$

and

$$b''_n = \frac{2}{N}\frac{\pi}{4}\sum_{k=0}^{N-1} x(k)S_o(\frac{2\pi nk}{N}). \tag{22}$$

First these two arrays $\{a''_n\}$ and $\{b''_n\}$ may be evaluated for a range of values of n. These coefficients correspond to single square wave functions. The frequency range for which the spectral amplitudes are to be evaluated is from f_1 to f_2. The cosine and sine waves being the sum of single square waves of corresponding frequency and their harmonics, the coefficients a'_n and b'_n come out as

$$a'_n = \sum_{m=0}^{\infty} \frac{-1^m}{2m+1} a''_{(2m+1)n} \tag{23}$$

and

$$b'_n = b''_n - \sum_{m=1}^{\infty} \frac{-1^m}{2m+1} b''_{(2m+1)n} \tag{24}$$

where m is an integer such that it is not a multiple of 3. These values of a_n and b_n correspond to the sum of the above sequences when infinite terms are taken into account. But since the signal is band limited, most of the higher terms come out to be zero as already explained earlier. One step of a''_n corresponds to a frequency interval $\Delta f = \frac{f_s}{N}$. The lowest frequency of desired spectrum, f_1 corresponds to a value of n equal to $\frac{f_1 N}{f_s}$ and the highest frequency f_h corresponds to a value of n equal to $\frac{f_h N}{f_s}$. This value of n may be termed as n_{max}. Thus, the sequences given above may be truncated to a few harmonics only as all terms a''_{kn} and b''_{kn} corresponding to values of $kn > n_{max}$ may be neglected.Consider, a signal $x(t)$ having a highest frequency component of f_m Hz. It is desired to analyze the signal in the frequency range f_1Hz to f_2Hz only. This is applicable to musical sounds. Let us consider a case where the desired spectral range is from 100Hz to 500Hz. The signal may contain higher harmonic contents, which are as high as 15KHz. A suitable filter may be used to attenuate the higher harmonic components lying above 1KHz. The frequency band of 100Hz to 1000Hz produces images at $\frac{1}{11}^{th}$, $\frac{1}{13}^{th}$, $\frac{1}{17}^{th}$ and higher fractional harmonics if only four terms are considered in a composite square wave. Thus, the maximum value of frequency at which the image distortion appears is $\frac{1000}{11}$Hz that is 91Hz. This does not disturb the range of 100Hz to 500Hz. The frequencies above 1000Hz are filtered out to achieve the exact spectral analysis of the signal using square waves in the desired range of 100Hz to 500Hz which is specified above. A signal consisting two sinusoidal frequency components of equal amplitude at frequencies 240Hz and 270Hz, is analyzed using composite square wave. The spectrum obtained is given in fig. 1 which shows no trace of any spurious frequencies.

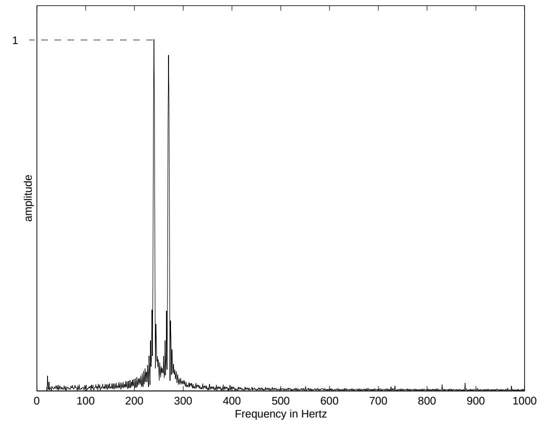

Figure 1. Frequency Estimation Using Composite Square Wave

7. ELIMINATION OF MULTIPLICATIONS IN COMPOSITE SQUARE WAVE FFT

For an N-point FFT, the discrete time signal is multiplied by $e^{\frac{j2\pi mk}{N}}$ where k represents the discrete frequency and m is the sample number. This factor is a sum of real and imaginary components given by $\cos(\frac{2\pi mk}{N})$ and $\sin(\frac{2\pi mk}{N})$. Each of these can be replaced by even and odd square waves. These square waves may be represented by $\mathrm{S_e}(k,m)$ and $\mathrm{S_o}(k,m)$ as,

$$\mathrm{S_e}(k,m) = \mathrm{sgn}[\cos(2\pi mk/N)] \tag{25}$$

$$\mathrm{S_o}(k,m) = \mathrm{sgn}[\sin(2\pi mk/N)]. \tag{26}$$

$\mathrm{S_e}(k,m)$ has a total of k cycles in the entire width of N samples. Thus, k represents the frequency of square wave if N samples cover a duration of 1 sec. It completes one cycle in N/k samples. The function $\mathrm{S_e}(k,m)$ has a value of $+1$ whenever $\cos(2\pi mk/N) \geq 0$ and -1 elsewhere. These square waves may be used in place of sine and cosine functions to obtain a transformation into frequency domain . This is similar to Short Time Fourier Transform (STFT) and may be termed as Fractional Fourier Transform (FrFT). These transforms result in even and odd functions as $X_E(k)$ and $X_O(k)$, which are given by

$$X_E(k) = \sum_{m=0}^{N-1} x(m).\mathrm{S_e}(k,m) \tag{27}$$

and

$$X_O(k) = \sum_{m=0}^{N-1} x(m).\mathrm{S_o}(k,m). \tag{28}$$

The overall value of $X(k)$ is given by

$$X(k) = X_E(k) - jX_O(k). \tag{29}$$

The function $X(k)$ contains the amplitude of the k^{th} frequency component plus certain fractions of its odd harmonics. The number of multiplications and additions come out to be equal to $2N^2$. But, because $\mathrm{S_e}(k,m)$ and $\mathrm{S_o}(k,m)$ have values +1 and -1 only, all the multiplications can be avoided. This is shown in the following analysis.

7.1. Signal Preconditioning for Reduced Computational Complexity

When the signal $x(m)$ is multiplied by a square wave $\mathrm{S_e}(k,m)$ or $\mathrm{S_o}(k,m)$ consisting of positive and negative half cycles of amplitude +1 and -1, and is then integrated, all the sample values coming within positive half cycle of the square wave get added up while those coinciding with the negative half cycle are subtracted.Let us consider that the number of samples in one cycle of square wave is $2l$. That is,

$$l = \frac{N}{2k}. \tag{30}$$

This can be viewed in fig. **??**. The function $X_E(k)$ can be written as sum of sequences, given by

$$\begin{aligned} X_E(k) &= \textstyle\sum_{m=0}^{N} x(m)\mathrm{S_e}(k,m) \\ &= \textstyle\sum_{m=0}^{\frac{l}{2}-1} x(m) - \sum_{m=\frac{l}{2}}^{\frac{3l}{2}-1} x(m) + \sum_{m=\frac{3l}{2}}^{\frac{5l}{2}-1} x(m) - \ldots. \end{aligned} \tag{31}$$

For each value of k, this summation has to be done for the entire sequence. The repeated addition can be avoided if a new sequence is initially obtained by integrating the sequence $x(m)$. Let this integrated sequence be $x_i(l)$, which may be defined as,

$$x_i(l) = \sum_{m=0}^{l-1} x(m) \tag{32}$$

and

$$x_i(0) = 0. \tag{33}$$

This reduces the summation of the sequence to simple subtraction as

$$\sum_{l_1}^{l_2} x(m) = x_i(l_2) - x_i(l_1). \tag{34}$$

Using this property, we may write an expression for $X_E(k)$ of eqn. 31 as

$$\begin{aligned} X_E(k) &= (x_i(\tfrac{l}{2}) - x_i(0)) - (x_i(\tfrac{3l}{2}) - x_i(\tfrac{l}{2})) + \dots \\ &= -x_i(0) + 2x_i(\tfrac{l}{2}) - 2x_i(\tfrac{3l}{2}) - \dots . \end{aligned} \tag{35}$$

Similarly, the expression for $X_I(k)$ can be obtained from the sequence $x_i(m)$ as

$$X_O(k) = -x_i(0) + 2x_i(l) - 2x_i(2l) + 2x_i(3l) - \dots . \tag{36}$$

Here, only additions and subtractions are involved and no multiplication is required. There are $2k$ additions or subtractions required to evaluate each $X_E(k)$ and $X_O(k)$. As k has to be varied from 1 to $\frac{N}{2}$ to evaluate all the coefficients, the total number of computations come out to be about $\frac{N^2}{4}$. Though the order of computations is N^2, the computations are simple addition of real values as compared to complex multiplications and additions required in FFT. Secondly, we need not evaluate all the coefficients $X(k)$ for $k = 1 \; to \; k = \frac{N}{2}$, but can limit our computation to desired frequencies only.

8. COMPUTATIONAL COMPLEXITY

A comparison of computational effort for the analysis of an audio signal using FrFT and standard FFT using SWSTT is made in the following analysis. The signal is passed through a low pass filter with a cut off frequency of 1000Hz before sampling.The range of spectral estimation is from $f_1 = 100$Hz to $f_2 = 500$Hz and the highest frequency in the signal $f_h = 1000$Hz. For FFT analysis, the number of samples should be a power of 2. When the number does not match, zeros are added to make the number equal to the power of 2. Here, let the sampling be done at 2048 samples per second and consider bins of 1 second. The computational requirements for this analysis will be of the order of $Nlog_2N$ for each bin. This figure comes out as 22528 per bin. Thus, the total number of basic add operations required for one complex operation is approximately 100. The total number of basic add operations for FFT analysis of a packet comes out as 2.5×10^6. The computational requirements of FrFT is evaluated below. Consider a square wave of 1sec. duration at a frequency of fHz. The signal $x(n)$ is first integrated to realize a new sequence $x_i(n)$. The number of half cycles in it are 2f. The evaluation of the even component at this frequency requires 2f additions or subtractions. The evaluation of odd component requires another 2f computations. Thus, the total number of computations at a frequency f come out as $4f$. When square wave harmonics upto seventh order are taken into account, computations at frequency values of $3f, \; 5f$ and $7f$ are also required. This requires a total of $\{f + 3f + 5f + 7f\} \times 4 = 64f$ computations. But many of these computations are not required. Firstly, the computations for all the frequency values at interval of 1Hz need not be computed. The number of computations depends upon the resolution desired. For musical pitch detection, 1% frequency resolution is adequate as estimated in section 2. Thus, values of a_n and b_n for frequencies 100Hz to 200Hz are computed with an interval of 1Hz, from 200Hz to 300Hz with interval of 2Hz and from 300Hz to 400Hz with intervals of 3Hz and so on. Secondly, the computations need to be carried out only for $3^{rd}, \; 5^{th}$ and 7^{th} harmonics lying in the desired range. The repetition of computations occuring can be avoided. For example, the analysis for 100Hz frequency requires values corresponding to 300Hz, 500Hz and 700Hz. These values need not be computed again for the corresponding frequencies. Also the coefficients corresponding to higher than maximum frequency in the signal come out to be zero. Thus, a''_n and b''_n need not be calculated for square wave frequencies above f_h. With the above modifications, the number of computations required for the analysis in the range 100Hz to 500Hz come out as 2.46×10^5 only. This number is less than the number of computations estimated for FFT. The above two schemes were implemented on a Pentium - III computer. The timings observed for both the techniques FFT and FrFT came out to be 9107 μsecs and 4013 μsecs respectively. This shows that FrFT proves faster than FFT. The important advantage of the FrFT is in the reduction of multiplications which makes the hardware realization more economical.

9. COMPOSITE SQUARE WAVE FOURIER TRANSFORMS APPLIED TO MUSIC SIGNALS

For a comparison of Composite Square Wave Fourier Transforms with FFT, these techniques are applied to a vocal music signal and the results are compared. A musical piece in Raga Yaman performed by Dr.S.B.Sharma, former Head, Dept. of Music, Dayalbagh Educational Institute, Dayalbagh, Agra, INDIA is selected. This raga has all the swaras of the octave in a sequence Sa, Re, Ga, Ma, Pa, Dha, Ni,$\dot{S}a$. $\dot{S}a$ denotes 'Sa' in next higher octave. A Tanpura is used as a drone to accompany the vocal sound. The duration of the sample is about 15 seconds. The signal is sampled at the rate of 48000 samples per second using a standard Sound Card. The signal is divided into 20 bins each of 32768 samples. Each bin occupies a time of nearly $\frac{2}{3}^{rd}$ of a second and the number of samples in each bin is chosen to be a power of 2. A digital filter algorithm is used to cutoff unwanted higher frequencies. The desired range of the first octave being 240Hz to 480 Hz, the range of search of the pitch is limited to 100 Hz to 720 Hz. The Pass band limit of this low pass filter is kept at 720 Hz. The pass band starts at 1KHz. Since, now the signal contains only the frequencies upto 1KHz, the signal is downsampled by a factor of 20. Thus, the sampling rate of the signal has been compressed to 2400. This bandlimited signal is subdivided into bins of 2048 samples each. The spectral analysis of each bin is done separately using all the three methods. The algorithms are coded in 'C' language and compiled on a UNIX operating system. The details of analysis of each method are given below. The two

Note	Ratio w.r.t note 'Sa'	Expected frequency (Hz)	Observed frequency(Hz)	
			FFT	FrFT
Sa	1	314.00	314	315
Re	$\frac{9}{8}$	353.25	353	353
Ga	$\frac{10}{8}$	392.50	394	394
Ma	$\frac{4}{3}$	418.66	415	415
Pa	$\frac{3}{2}$	471.00	468	468
Dha	$\frac{27}{16}$	529.87	522	523
Ni	$\frac{15}{8}$	588.75	586	586
$\dot{S}a$	2	628.00	625	625

Table 1. Comparison of theoretical estimates and observed frequencies

different tools i.e. the standard FFT and the proposed FrFT showed similar results as shown in table 1. The FrFT is matching with FFT closely. Some other frequency components are found to be present throughout the sample duration. These are sub-harmonics of the voiced sound and frequencies sounded by Tanpura. The voice signal is complex and has several harmonics and sub-harmonics. In-spite of these components, the true tone is clearly visible. The musician sounds one note and changes to next note in a sequence. The voice slightly dips while this change occurs. The true frequency is observed nearly at the center of the time taken by a note. The results obtained by FFT and FrFT are matching exactly. The timings of the two methods are

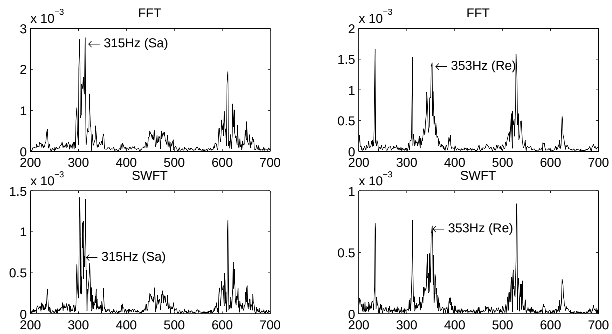

Figure 2. Notes 'Sa' and 'Re' as obtained from FFT, FrFT

also compared for a 2048 point analysis and 32768 point analysis. The time taken for analysis is given in table. 2. The above data shows that the FrFT consumes the smaller time duration as

Tool	Time (ms) (N=2048)	Time(ms) (N=32768)
FFT	9.1	799.6
FrFT	4.0	253.2

Table 2. Timing comparison for both techniques

compared to FFT.

10. CONCLUSIONS

The proposed FrFT extract exact Fourier Coefficients of a band-limited periodic signal for a limited range of frequencies by using square waves in place of sinusoidal waves. This results in simpler arithmetic as the need of multiplication is avoided. Though the number of computations is higher, both the computational time and hardware requirements are reduced.

REFERENCES

[1] J.Pender and D.Covey, "New Square Wave Transform for Digital signal Processing", *IEEE Transactions on Signal Processing* **25**, vol.41, March 1992.

[2] K.U. IngardM.P.Lamoureux, "The Poorman's Transform: Approximating the fourier Transform Without Multiplication", *IEEE Transactions on signal Processing* **25**, vol.41, March 1993.

[3] J.D.B.Nelson, "On the Co-efficient Quantization of the Fourier Basis", *IEEE Transactions on Signal Processing* **25**, vol. 51, July 2003.

[4] Y.C.Lim and B.Farhang-Baroujeny, "Fast Filter Banks", *IEEE Transactions on Circuits and Systems-II, Analog and Digital Signal Processing,* **25**, vol.39, May1992.

[5] S.S.Bhave, "Letter recieved through private communication."On the theroy and design of acoustic resonators", *CSC Group* **25**, TIFR, Mumbai.

[6] D.O'Shaughnessy, *Speech communications, Human and Machine*, IEEE Press, 2000

[7] K.C.Beauchamp, *Walsh Functions and Their Applications*, New York: Academic, 1975.

[8] K.C.Beauchamp, *Applications of Walsh and Related Functions*, New York: Academic, 1984.

[9] A.W. Leissa, *Vibration of shells*, American Institute of Physics, Woodbury, New York, 1993.

[10] Charu and V.G.Das, "Square Wave Technique for Estimation of Pitch in Music Signals", *Proceedings of International Congress on Acoustics 2004*, Kyoto, Japan, 4-7 April 2004,