APPENDIX A

Frequency Estimation

A number of techniques are available for estimating the frequency of digitally recorded sound. Some of the techniques used in this study are briefly described below:

The commonest technique is the Discrete Fourier Transform (DFT), which is based on the Fourier theorem that every periodic signal can be represented as a sum of one or more harmonics of the fundamental with suitable weights. The analysis thus shows what frequencies are present in the sound, and can also show the spectrum (or the composition of the harmonics of the sound). When dealing with data sampled over a short duration, this technique creates certain problems due to the sudden cutting off (truncation) of the waveform at the two ends. This is partly taken care of by the process of 'windowing', i.e., modifying the sample so that its amplitudes taper at the ends. Further, when the sample analysed does not cover the exact number of periods of the sound (which is generally the case), the frequencies (the peaks in the spectrum) do not really correspond to the actual frequencies, and interpolation techniques have to be used to estimate the frequency.

The accuracy of the estimate improves with increase in sample length (i.e., in samples of longer duration), but, while analysing gamakams, excessively long samples can miss the fine changes of frequencies. In spite of these drawbacks, very fast computations can be done with the DFT technique by using Fast Fourier Transform (FFT), which requires the number of samples to be a power of 2: such as 128, 256, 512, etc. Because of the high speed of calculations, it is possible to calculate the frequencies at short intervals (such as a few milliseconds) of a sample and use them to plot a graph showing the movement of the melody. This is very useful for a general view of the gamakam, and also enables further detailed measurements at the required points.

Autocorrelation is another technique which helps us to match a select number of digital samples with the same number of samples shifted to different points in time. In a perfectly periodic wave, there will be a hundred per cent match between two groups of digital samples separated by one period (reciprocal of the frequency), two periods, etc. A similar technique requiring less computation is Average Magnitude Difference, which also involves matching two sets of samples by measuring the absolute value of the differences between the amplitudes of the corresponding points of the two sets.

A third approach is 'zero crossing'. Here, the two adjacent points at which the amplitude of the wave crosses zero in a particular direction are identified, and the time difference between the two points is taken as one period. This can work correctly only if there are no harmonics or the harmonics are weak, or where the original sound is filtered to remove higher frequencies.

In the present study, the application of this concept was further extended. The periodicity of a waveform is easily recognizable when depicted visually as a graph. The program written (which I can supply to users) enables blocking a certain number of periods (between two maxima or minima or two points of zero crossing), and calculating the frequency after automatically making corrections for locating the correct maximum, minimum, or zero crossing, and also interpolating to get a more accurate point (as we are dealing with discrete samples).

The last technique comes nearest to the techniques used earlier by C. S. Ayyar and Jairazbhoy and Stone, and can also be taken as the one closest to the pitch felt by the listener, especially when the fundamental is strong. (It can, however, be argued that the overall pitch sensation is based on all the harmonics, and when the overtones are not perfect multiples of the fundamental but quite strong, filtering can distort the results. This problem is discussed later in this appendix.)