

ISIS, AN ALTERNATIVE APPROACH TO SOUND WAVES

Clarence Barlow
Royal Conservatoire
Juliana van Stolberglaan 1
2595 CA The Hague
The Netherlands
18 June 2005

ABSTRACT

ISIS, for 'Intra-Samplar Interpolating Sinusoids', is a means of mathematically interpolating sine wave segments between the samples of a sound wave recording (the word "sample" is here used as in "sample rate"). The sound wave is thus representable as a sequence of sine-tone pitches in extremely small time windows in the microsecond domain, rendering the wave audible as a rapid sine-tone 'melody' e.g. by slowing it down 4000 times, a feature given neither by regular time-stretching methods nor by cubic spline-interpolations. Conversely, this ISIS-extracted sequence can be accelerated to regain the original sound; indeed any melody can be similarly sped up to form a sound wave. The main application of ISIS is at present in the technical unification – for compositional purposes – of the micro- and the macro-temporal domains (pitch, timbre and rhythm) into one comprehensive field.

Methodologically, ISIS comprises the three areas *Analysis*, *Morphosis* and *Synthesis*, described below after the following historical account ('Morphosis' is a biological term for the "sequence of development or change in an organism or in any of its parts").

1. BACKGROUND

The beginnings of ISIS go back to 1969, when I first started to doubt the common adage that white noise was the "sum of all frequencies". This had implied that since any noise band, no matter how narrow, would necessarily contain an infinite number of simultaneous frequencies, each of these frequencies would as a consequence have to have zero amplitude, a model I did not find useful.

In 1971, after starting to work with stochastic computer programming techniques, I began to savour the idea that white (or indeed any type of) noise could in fact be regarded as a single sine wave of rapidly changing frequency. I expected that the probability of this hypothetical sine having a certain frequency at any moment would relate in some simple way to the spectral amplitude of that frequency in the noise.

Soon after the commercial introduction of MIDI in 1984, I sent rapid (upto 200 note-per-second) random-pitched MIDI-note streams to an FM-synthesizer generating sine tones and other sounds including that of a simulated piano: the notes' pitch probability was proportionally equated to the spectral amplitudes of the corresponding frequencies of well-known timbres such as vowel phonemes. In all cases, the results bore out my hypothesis: note streams generated by these phonemes, melodies when under 20 nps (notes per second), blurred at higher rates to pitch clouds bearing the phonemes' respective timbres. I called this method *Spectastics*, from 'Spectrally defined Stochastics'.

In the 1990s I used spectastics in two pieces, one (involving a live player piano) in 1994, and again in a piano solo of 1998. Yet, though MIDI can in theory transmit upto 600 nps, my synthesizer couldn't handle more than 200 and I had to limit myself to only 50 and 11 nps respectively in the two above-mentioned pieces for evident mechanical and biological reasons.

It was not until 2001 that I finally set myself the following consequential proposition: assuming a note speed going up into thousands per second, could a situation finally arise where each sound wave sample results from a single input note? In reverse, I asked myself if tangible pitch information could be extracted from a single given pair of contiguous samples. And indeed: with recorded samples seen as points from left to right in a space bounded vertically by ± 1 , a sine wave segment connecting any pair of adjacent points (and – for histogrammatic reasons explained later – tangent once each to its ± 1 y-values) can be shown to possess a unique frequency near to the sampling rate. Not only is any sound wave thus interpretable as a pitch sequence running at sample speed, one can also reconstruct a wave acoustically indistinguishable from the original by applying the reverse procedure to the obtained sine pitches. I named this set of procedures *ISIS*. It is currently being developed in Max and Super-Collider at the Royal Conservatoire The Hague. Sound examples are available for each section of the following text and for the explanatory diagrammes.

2. ANALYSIS

ISIS assumes the amplitude of all interpolating sines to be constant and at maximum, i.e. the sine segment is tangent to $y = +1$ and to $y = -1$ once each between two adjacent samples. The final phase of any sine segment and the initial phase of the one following are identical at the connecting sample. This method yields solutions according to the simple formula

$$I = \log \left[\frac{\{\arcsin(S_2) - \arcsin(S_1)\} / 2\pi}{\log(2)} \right] \quad (1)$$

where S_1 and S_2 are two successive samples (bounded by $y = \pm 1$) and I is their ISIS pitch interval expressed as a deviation in octaves from the sampling rate, an arbitrary value which is unnecessary in the formula but must be the same in synthesis to successfully reconstruct the analysed wave. The ISIS-extracted frequency implied by the interval I is of the serial form $\mathcal{Q}(n+2^I)$, where n is a whole number and \mathcal{Q} the sampling rate, and – by appropriately setting n – is the frequency closest to \mathcal{Q} (for reasons of histogrammatic concision; this is why the sine is once tangent to ± 1).

Figure 1 shows an example of a set of randomly chosen samples (seven dots connected by a faint line – the “sound wave”) analysed according to ISIS: each pair of samples is transected by a segment approaching one period of a sine curve, the frequency of which is given at the top of the diagramme. The phase of the sine segments meeting at the samples is also shown in degrees (in the range -179° to $+180^\circ$).

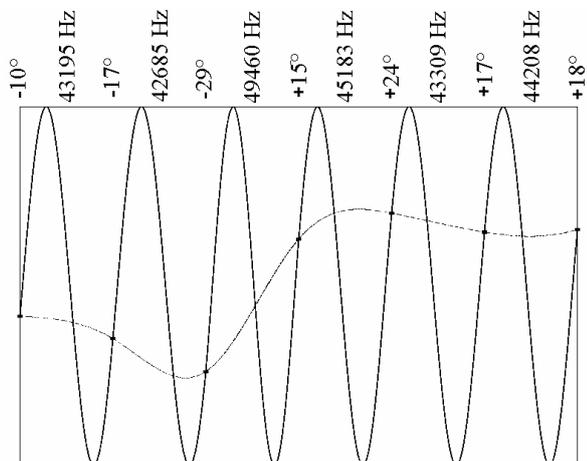


Figure 1. The ISIS principle: interpolating conceivable sine segments of determinable frequency between samples of a sound wave.

The fact that with a sampling rate of 44100 Hz these pitches mostly transcend the range of human hearing does not prevent their transposition to lower regions for musical purposes as will be evident below.

For ISIS analysis, I wrote a computer program termed *AnalISIS*, which accepts as input a sound wave in a standard file format. The output file, in so-called SIS (‘Sustainable Interval Sequence’) format, is a list

of ISIS intervals in cents coupled with their ‘sustenance count’, the number of interpolations for which the given interval is to be extended, factually the duration continuance of the interval in samples. For example, the values ‘-3.21 4’ mean 3.21 cents below the sampling rate for a total of 4 sine-segments (5 samples). Since the phasing is always contiguous, a continuous smooth sine is created all the way through the specified duration.

Figure 2 shows the ISIS-treatment of a tape-snippet from an interview I made with Morton Subotnick in December 1999, with him saying at this point “..left alone at the turn of the century..”, a comment on avantgarde composers then in terms of their erstwhile followers. Here the SIS pitch extraction of the original wave is supplied in tabular form at left (a short excerpt from the beginning) and also in graphical form at top right, as well as in a neat SIS-histogramme at lower right, of which the total pitch range was found to be roughly ± 150 cents; notice how the SIS interval (pitch deviation) values closely follow the amplitude of the source wave, shown reconstructed at centre right and graphically and acoustically indistinguishable from the original.

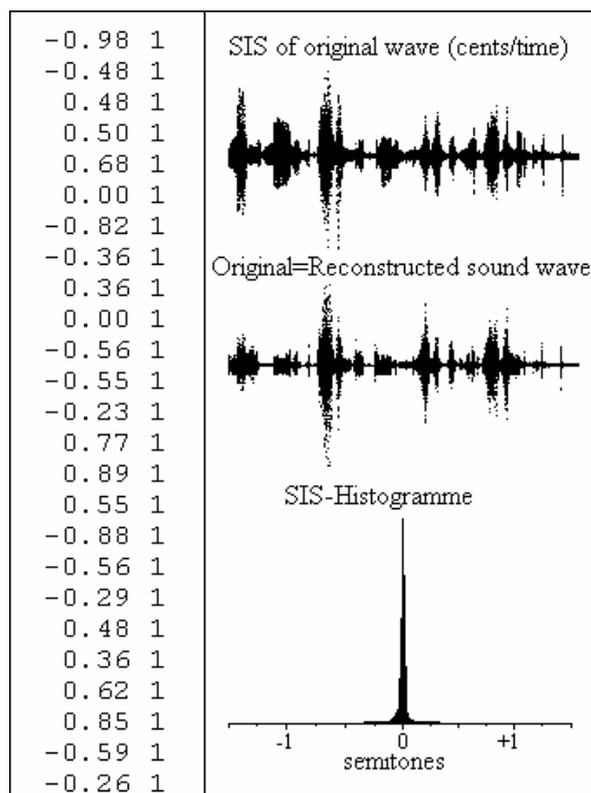


Figure 2. Left: start of Sustained-Interval-Sequence (SIS) file extracted from a sound wave; values in cents and sustenance (here 1), Top right: graphic depiction of this SIS file, Centre right: sound wave (ISIS-reconstructed and indistinguishable from original), Lower right: histogramme of SIS-file values.

3. MORPHOSIS

Having extracted a SIS cents file from a sound wave, one can now perform various pitch-transformations on it. In the main, these are operations I call Stretching, Shifting, Sustaining, Skipping, Smoothing etc., shown below.

3.1. Stretching

This involves multiplying the SIS cent values by a certain factor. Figure 3 shows the resynthesized waves after respectively doubling and halving the SIS-intervals of Figure 2. Apart from a marked change in the RMS amplitude, one notices a striking DC-bias; and though it seems as if the waves to distort, this is not so: all samples are computed sine functions and therefore within the prescribed ± 1 limits.

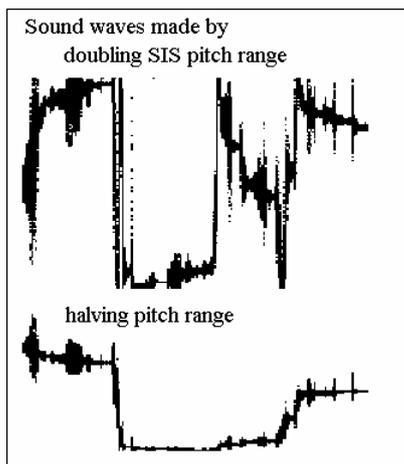


Figure 3. Sound waves obtained by the doubling (above) and halving (below) of the SIS pitch values shown in Figure 2.

3.2. Shifting

Here the contents of the SIS file are raised or lowered by a certain pitch interval. Interestingly, and predicted by the synthesis algorithm, an intrinsic sine tone appears of maximum amplitude and frequency

$$\mathcal{Q}(2^I - 1), \quad (2)$$

where \mathcal{Q} is the sample rate and I the interval in octaves; removing this tone leaves the resynthesis still altered, but closer to the original. Figure 4 shows the ISIS-shifted sound wave with the intrinsic tone (above) and without it (below). A downwards shift by say 7 octaves brings the ISIS-melody (SIS) from typically 44100 Hz down to 345 Hz, well within the audible pitch range.

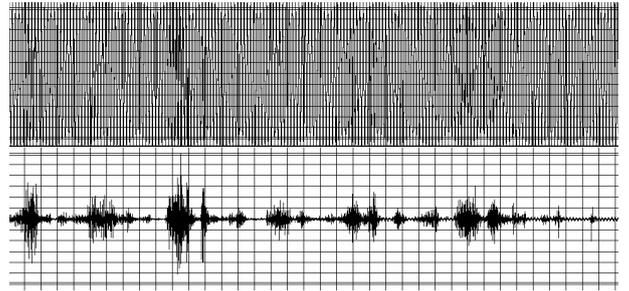


Figure 4. Resynthesized ISIS-shifted sound wave with intrinsic tone (above) and with the tone removed (below).

3.3. Sustaining

Here the interval sustenance count found in the analysis is changed during resynthesis. For instance, changing all values e.g. from 1 to 2 will cause the sound to drop an octave while taking twice as long. A change e.g. to 4900 causes the individual notes of the SIS to drop in speed from 44100 nps to only 9 nps (= 44100/4900), well within the audible tempo range.

3.4. Skipping

In the case given above (sustaining by 4900 or by any number larger than 1) the output will be longer in time than the input (e.g. 4900 times longer), unless sufficient input samples are skipped during analysis. Equating the sustain and the skip values will enable real-time work: the output runs at the same speed as the input. If skip exceeds sustain, the sound result will be faster.

3.5. Smoothing

This means tempering – cents are ‘smoothed’ into steps, multiples of a given value, set in Figure 5 at one and ten cents, respectively. Here, too, the DC-bias varies and the sound deteriorates strongly with larger steps.

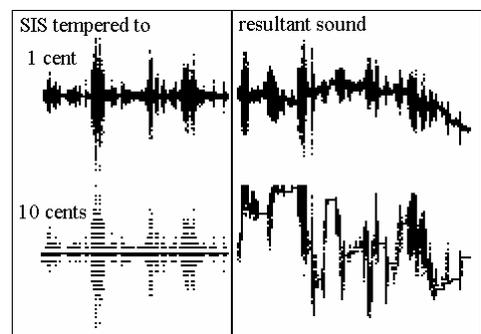


Figure 5. Pitch-time diagrammes of SIS intervals smoothed to steps of 1 and 10 cents (left); resultant sound (right).

4. SYNTHESIS

The process of ISIS pitch extraction is reversible by inputting an SIS sequence into the reverse algorithm

$$S_2 = \sin\{\arcsin(S_1) + 2\pi \cdot 2^I\}, \quad (3)$$

where I is the ISIS pitch deviation in octaves from the sampling rate and S_1 and S_2 any two successive samples (within ± 1). S_1 can be initially set to 0.

The pitch input can be of various types, of which two categories will be described, Acquired and Algorithmic.

4.1. Acquired Input

An example is the flowing 7-bar melody from J.S.Bach's *Jesu Joy of Man's Desiring*. These 63 notes, looped 700 times, give 44100 pitches forming at sampling speed a sound wave (seen vertically at left in Figure 6) lasting one second; the SIS source is at far left (Middle E is taken at 0 cents) and the sound's striking FFT at right.

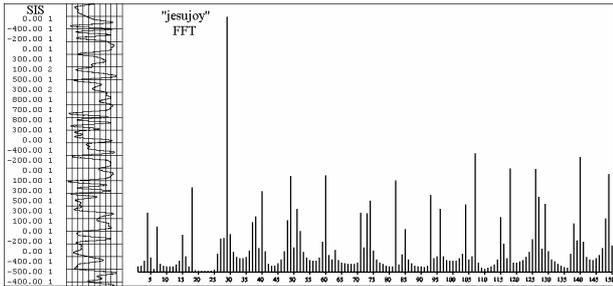


Figure 6. A melody (“Jesu Joy..”) as a SIS (see excerpt far left) for ISISynthesizing a sound wave (excerpt left), the FFT of which is on the right.

In stark contrast is the complete 36-bar melody of Ravel's *Bolero*, sped up 3000 times and taking 29 milliseconds; the start of the SIS file and sound wave as well as the sound's FFT are to be seen in Figure 7. Here it was Middle D which was set at 0 cents.

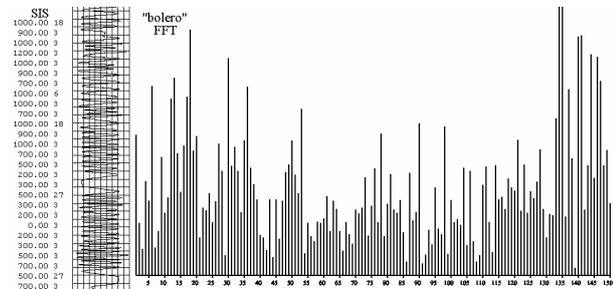


Figure 7. The melody of Ravel's “Bolero” as a SIS (excerpt far left) for ISISynthesizing a sound wave (excerpt left), the FFT of which is at right.

4.2. Algorithmic Input

Two input types are outlined here: Deterministic and Spectastic.

4.2.1. Deterministic input

Sped up sufficiently, an octave-tremolo of sinusoids at 440 and 880 Hz merges into a near-sine at 660 Hz. In the curve at left in Figure 8, the steeper intra-samplar sine segments are at 880 Hz, the less steep at 440 Hz, together forming the arithmetic mean 660 Hz, shown over a longer time period at centre right with the FFT at lower right and the SIS-histogramme at top right.

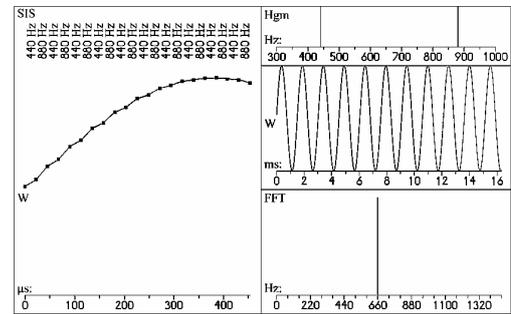


Figure 8. Sine-tremolo (see histogramme top right) sped up to sampling speed, resulting in a near-sine (see segments at left as well as the curve and its FFT at right).

4.2.1. Spectastic input

FFTs of the phonemes [u], [a] and [i] yield spectastic streams in which the pitches' probability equals their spectral amplitude (see “Background” above). When accelerated, the note-streams form pitch-clouds timbrally resembling the phonemes in question at 200 nps but less so at sample speed – see Figure 9: the narrow [u] spread, both in its SIS-histogramme and its FFT, getting wider through [a] to [i] is apparent.

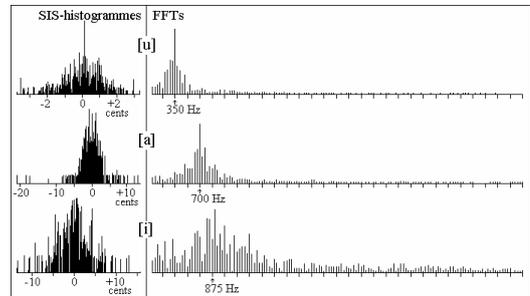


Figure 9. Three phoneme-generated spectastic streams at sample speed shown as SIS-histogrammes and as FFTs