Audio Visualization in Phase Space *

David Gerhard
School of Computing Science
Simon Fraser University
Burnaby, BC V5A 1S6, Canada
E-mail: dbg@cs.sfu.ca

Abstract

There are several modern methods to visualize sounds, from oscilloscope and spectrometer to colour organs and strobe lights. Phase space is a relatively new way to visualize sounds. Originally used to better understand the chaos of strange attractors and other non-periodic systems, it can be used to observe the regularity of periodic and pseudo-periodic signals. Besides extracting information and interesting complexity of sound signals, these phase space images are attractive and compelling to look at.

1 Introduction

The analysis of sound has been a topic of interest to researchers since the early 1600s, when Galileo and Mersenne did experiments to relate the pitch of a sound to its frequency. Sound visualization has been used as a method of sound analysis since the late 1700s, when Higgens observed the effect of a sound on a candle flame. Later, Rudolf Koening built an instrument for observing the flame of a candle over time, and this instrument generated images called "manometric flames." There is a good introduction to the history of sound visualization in [8].

Humans have been interested in observing sound signals because we are (or we think we are) "better" at seeing than hearing. There are more neurons in our brains devoted to seeing than to hearing, and we think we are more able to analyze complex data when it is in the visual domain. Perhaps it is because we feel we will be more able to understand something if we get information about it using two of our senses instead of one. Whatever the reason, using a non-traditional sense is an interesting way to "see" a problem in a new form, from a new point of view.

The oscilloscope is currently the tool of choice for observing sound signals (and in fact any electrical signal) in the visual domain. It takes as input a voltage level and provides as output a trace of the voltage level over a small interval of time. Two voltage levels from two different sources can be displayed, one against the other, to see the correlation between these signals.

*This research is supported in part by The Science Council of British Columbia

The oscilloscope allows us to observe, for example, the pleasing and smooth shape of the sinusoidal acoustic signal, which we understand to be the preferred basis for all other signals. It allows us to observe, for example, that a square wave can never quite be generated as an acoustic or electric signal, a fact that our ears alone would not be able to deduce. As well, the oscilloscope allows us to see interesting and somewhat visually pleasing shapes corresponding to signals that are not traditionally or naturally visible.

There have been other interesting attempts to visualize sound, some with the mandate to entertain rather than understand. A common example is the colour organ, which provides voltage to a number of different lights of different colours, depending on the energy in a particular frequency band of the signal. This device is mostly used in dance parties and discos to add visual stimulation to the already excessive audio stimulation.

Phase space is yet another way to visualize audio signals. It generates pleasing and interesting shapes, and it provides a level of insight into the nature of sounds that is not available in the standard Fourier representation (which I will review briefly) or the representation we are most familiar with - the common auditory sensation of these signals.

2 Background: Signals, Fourier and Phase

A sound signal consists of varying levels of air pressure on a human tympanic membrane or the diaphragm of an electronic microphone. As such, the signal itself is one-dimensional: it can be plotted as pressure versus time. This is the representation that is seen on the cathode-ray tube of an oscilloscope. If the signal is periodic, we see a repeating stationary signal, but if the signal is not periodic, the image we see is constantly changing and appears chaotic.

Humans have the incredible ability to extract large amounts of information from this one-dimensional signal. The basilar membrane in the inner ear vibrates at different frequencies in different places: two simultaneous tones set up two different patterns of vibrations at two different locations on the basilar membrane. This process is similar to the mathematical construct of the Fourier Transform, which breaks the incoming signal into component sinusoidal functions, making the harmonic components of a signal apparent.

Each sinusoidal component in the Fourier representation of a signal has three associated values: frequency, amplitude, and phase. The components are commonly plotted as amplitude versus frequency, with phase versus frequency in another graph. The phase of the component corresponds to where in time the sinusoid "starts." Two sinusoids with the same frequency and phase will have their peaks and troughs line up exactly. Two sinusoids with the same frequency but opposite phase will have their peaks and troughs cancel each other out. The phase information is often ignored in a Fourier representation because the ear is "phase deaf," meaning that two stationary sinusoids of the same frequency but different phases will sound the same. Since sound waves are additive, any combination of sinusoids will sound the same as the same combination with different phases, provided the phase of any particular sinusoid does not change over time.

In a more general context, the phase of a signal at a time t corresponds to where in the cycle that signal is at t. A common way to represent this is to plot the signal against its first derivative, the thought being that the signal value and the value of the first derivative provide information about the state of that signal at t. This concept is discussed in a more general form and extended into higher dimensions in the next section.

3 Phase Space and Pseudo-Phase Space

Phase space is an n-dimensional space ϕ_n of vectors, created from a one-dimensional function $f(t)$, and is a subspace of $\mathbf{R}_n$. The tuples making up the space contain the value of the original function and the value if its first $n-1$ derivatives. The tuples are ordered according to the input variable t. For a three-dimensional phase space representation of $f(t)$, we write

$$\phi_3(f) = (f(t), f'(t), f''(t)) \tag{1}$$

and in the general n-dimensional case,

$$\phi_n(f) = (f(t), f'(t), \ldots, f^{(n-1)}(t)) \tag{2}$$

where $f^{(n)}(t)$ is the n-th derivative of $f(t)$.

The commonly used discrete form of this transform is called "pseudo-phase space," hereafter PPS, and is performed on a sampled signal $G[k]$. Tuples are created containing the value of the signal at k, as well as the value of the signal at k plus some delay τ. For a three-dimensional PPS $\Phi_n(G)$, we write

$$\Phi_3(G) = (G[k], G[k-\tau], G[k-2\tau]) \tag{3}$$

where τ is a suitable delay (in number of samples) between measurements. The general n-dimensional space is then

$$\Phi_n(G) = (G[k], G[k-\tau], \ldots, G[k-(n-1)\tau]) \tag{4}$$

3.1 Discrete Phase Space vs. Pseudo-Phase Space. It should be clear to mathematicians, engineers and anyone who knows calculus that PPS is not simply the discrete form of phase space. It is not a true phase space, and this is why the terminology "pseudo" is used to describe this space. Discrete phase space is a direct translation of continuous phase space into discrete phase space, where the derivatives are calculated using filters. The reason that PPS is used instead of discrete phase space is that the algorithm is much less computationally intensive, and the result is similar.

3.2 Periodic Signals in Pseudo-Phase Space. Figure 1 presents a two-dimensional PPS representation of broadband noise (more specifically, crinkling newspaper). For all of these PPS diagrams, the time-domain signal is in a rectangular window at the top of the diagram, and the two-dimensional PPS representation is in a square window at the bottom of the figure. I will

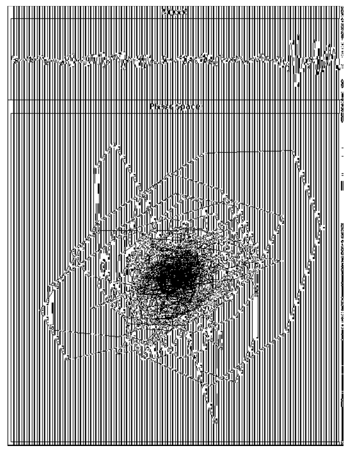

Figure 1: Phase space representation of broadband noise

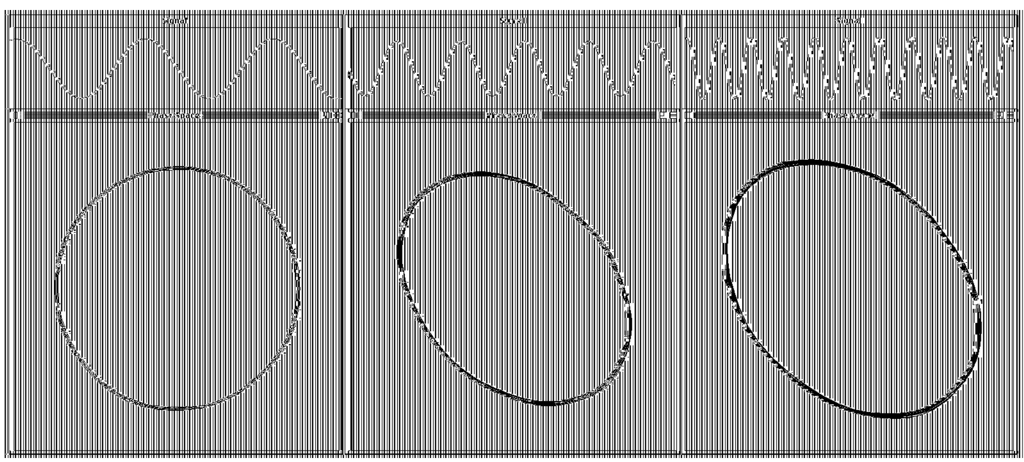

Figure 2: Phase space representation of sinusoidal signals. From left to right: 220 Hz, $\tau = 50$ samples; 440 Hz, $\tau = 20$ samples; 880 Hz, $\tau = 10$ samples. The sampling rate is 44,000 Hz.

restrict the first part of this discussion to two-dimensional PPS because it is easier to visualize, it is easier to present on paper (which is notoriously two-dimensional) and it is sufficient for most of our discussion.

In Figure 1, we see that there is not much obvious structure in two-dimensional PPS or in the one-dimensional time-domain representation. This is to be expected, because broadband noise has a high dimensionality - it takes a lot of different basis functions to be able to fully represent it. A periodic signal usually has a much lower dimensionality, depending on the harmonic content of the signal. The PPS representation of a few sinusoids of varying frequencies is displayed in Figure 2.

The PPS representation of a sinusoid is a circle or an ellipse, depending on the value of τ. The representation will be a perfect circle, as in the left-hand example in Figure 2, if τ is $\frac{1}{4}$ of the period of the sinusoid, plus some multiple k of the period [5]. In the case of the 220 Hz sinusoid, each period takes $\frac{1}{220} = 0.00455$ seconds, and at 44,000 samples per second, this is equivalent to 200 samples per period. One-fourth of 200 is 50, and it is clear in Figure 2 that $\tau = 50$ provides a circle.

In the two other cases presented in Figure 2, the figures are not exact circles because τ is not equal to one-fourth of the period. What is interesting is that, apart from the size of the ellipse, which depends only on the amplitude of the original signal, these figures are the same. In phase space as well as PPS, frequency has no bearing on the shape of the cycle. Two signals that are the same except for frequency can be made to look identical in PPS by matching amplitudes and matching the ratio of τ to the period of the signal.

What *does* change with frequency is the number of samples in each cycle in PPS. There are many more samples in each cycle of the 220 Hz sine wave, and therefore more samples in each pass around the circle in PPS. By measuring how many samples are required for a PPS cycle to repeat itself, it is possible to determine the frequency of the original signal. This is one way to extract information out of this representation. Section 4 describes more ways to extract information from the PPS.

3.3 The Effect of τ. The choice of τ for the PPS representation is important, especially as it relates to the period p of the (periodic) signal. It has already been shown that for a sinusoid, if $\tau = \frac{p}{4} + kp$, the PPS representation will be a circle. If τ is chosen to be $\frac{p}{2} + kp$, the resulting PPS representation will be a straight line along the line $y = x$, and if $\tau = kp$, the PPS will be a single point. These last two statements are true even if the original signal is not a sinusoid.

While any τ will provide a PPS representation with interesting information, one can choose τ carefully to create pleasing and even artistic PPS representations.

3.4 The Shapes of Different Signals. Sinusoids make circles in PPS, and circles are interesting, but are there more interesting signals, making beautiful string-art-type images? Well, there are! Figure 3 shows six different PPS representations of signals created from superimposing two sinusoids of different frequencies.

Notice in Figure 3 that the three most fundamental intervals, on which the equal-tempered scale of music is based, have relatively simple cyclic representations in PPS. These intervals are the perfect fourth, in Figure 3c, with a frequency ratio of $\frac{4}{3}$; the perfect fifth, in Figure 3d, with a frequency ratio of $\frac{3}{2}$; and the octave, in Figure 3f, with a frequency ratio of $\frac{2}{1}$. The equal-tempered scale that most modern western music uses is based on making these three intervals "true," or as close as possible to the ideal small-integer ratios. The other intervals are slightly mistuned from small-integer ratios, and this is the reason for the "mesh" effect in Figure 3a, b, and e. For more on musical scales, see [2] and [7].

These three intervals are easy to identify in PPS. An octave is a cycle with a "dent" in it, with the same period as the base note. There are two "lobes" in the PPS representation. The perfect fourth has seven lobes in PPS, with a period three times the period of the base note. Similarly, the perfect fifth has five lobes, and a period twice that of the original signal.

It should be noted that these are very simple signals, containing one or two sinusoids. Natural sounds have much more complex PPS representations, and the resulting patterns are very different and worthy of more study. Figure 4 presents three PPS representations of natural musical sounds.

It is clear that there is much underlying complexity in the pseudo-phase space representation of these signals, and one possible research direction is to visualize a higher-dimensional PPS representation.

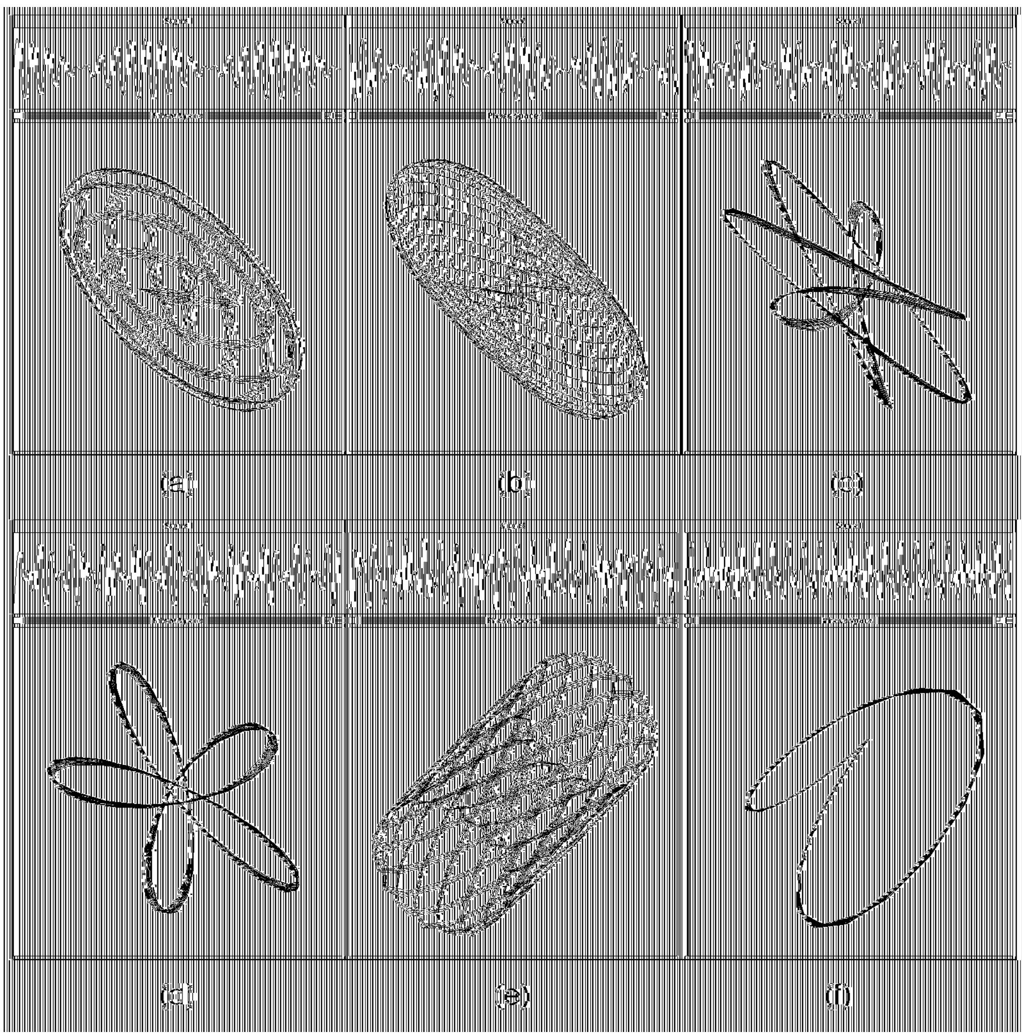

Figure 3: Phase space representations of two-note "chords" made by adding sinusoids of different frequencies. In all of these figures, $\tau = 20$ samples. a) Minor tone, A (1760 Hz) + B (1975.5 Hz). b) Minor third, A (1760 Hz) + C (2093 Hz). c) Perfect fourth, A (1760 Hz) + D (2349.3 Hz). d) Perfect fifth, A (1760 Hz) + E (2637 Hz). e) Minor seventh, A (1760 Hz) + G (3136 Hz). f) Octave, A (1760 Hz) + A (3520 Hz). The sampling rate is 44,000 Hz.

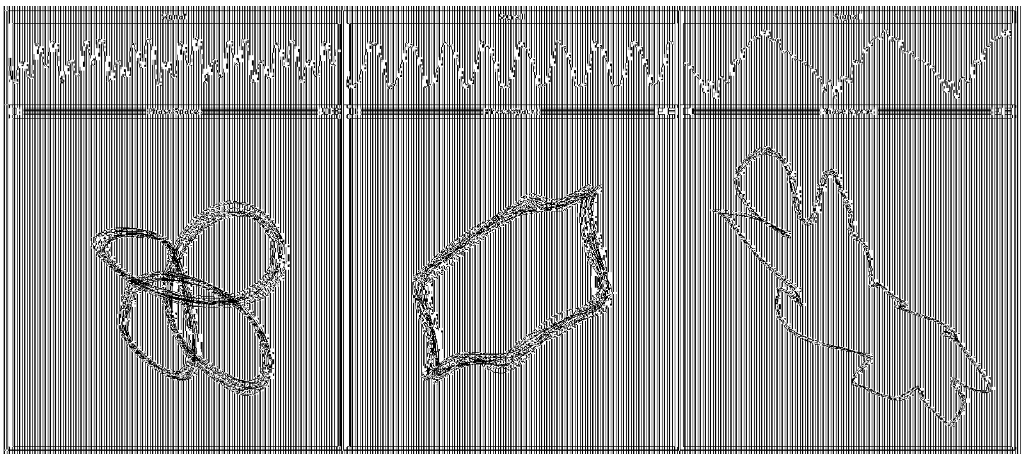

Figure 4: Phase space representations of single notes on various instruments. From left to right, these are an Accordion, a Penny Whistle, and a human voice singing the phoneme "Eee."

Three dimensions is not difficult to visualize in two dimensions (on paper) - the human mind is capable of resolving three axes. When the signal suggests a dimensionality of five or six, it is harder to visualize, and we must turn to such visualization tools as colour, motion, cutaway and the like.

4 Information Extraction

It might be asked at this point, "Sure, it's pretty, but does it *do* anything?" It seems as though there is much information that might be extracted from the PPS representation of a sound signal. If we would like to use the PPS to classify sounds, for example, we need to investigate what properties of the PPS remain consistent over a class of similar signals. If we wish to extract information about a particular signal, we need to present various signals to the PPS that have only a particular feature in common, and see if the PPS has some consistency across this set. Also, we need to make sure that the PPS does not have this consistency when the feature is *not* present.

4.1 Frequency Detection. It was stated earlier that a periodic signal will have a cyclic PPS representation that repeats itself once for every cycle of the original signal. This can be used to discover the frequency of the original signal. When the trace passes a specific point in the cycle, the time is recorded, and compared to the last time the trace passed that point. The difference in time is the period of the signal. This measure must be done many times and the results averaged, because PPS is a digital representation and as such is able to measure time at a resolution equal to the reciprocal of the sampling rate, and no better.

4.2 Interval Classification. We have seen that we can classify simple pairs of sinusoids by counting the "lobes" in the PPS. Five lobes implies two sinusoids of similar amplitude, with a frequency ratio between them of $\frac{3}{2}$. Similarly, seven lobes implies a frequency ratio of $\frac{4}{3}$.

The extraction of other information from the PPS has not been investigated in this work, but one could imagine using the PPS to classify the timbre of a signal, perhaps by investigating the fractal dimension of the cycle in PPS, or some other measure of complexity. The overall power of a signal could be extracted by finding the average distance of the trace from the origin, and doing some mathematical manipulations.

PPS is domain, like the time domain and like the spectral domain, where information can be extracted about signals. It has the advantage of being computationally efficient, but it does not seem to separate spectral information. Perhaps, with more study, the PPS representation will be able to provide more information.

5 Conclusions

Phase space is a relatively new and interesting way to visualize sound signals, by plotting the value of a signal against the values of its derivatives. Pseudo-Phase Space is a simplification of phase space, where the value of the signal is plotted against a time-delayed version of itself. Periodic signals create repeating cycles in phase space and PPS, sinusoidal signals produce circles and ellipses, and combinations of two sinusoids create interesting symmetric patterns. Natural musical signals with higher harmonics have intricate "curlicue" type cycles. The images that are produced in phase space are attractive line art, from the simply symmetric to the intriguingly chaotic. PPS can be used to extract information from a signal that might not be immediately obvious in the original time domain representation.

References

[1] H. Abarbanel, *Analysis of Observed Chaotic Data*, New York: Springer. 1996.

[2] D. Gerhard, *Automatic Interval Naming Using Relative Pitch*, Bridges: Mathematical Connections in Art, Music and Science, Winfield KS. August 1998.

[3] V. Gibiat, *Phase Space Representations of Acoustical Musical Signals*, Journal of Sound and Vibration, Vol. 123, No. 3, pp. 529-536. 1988.

[4] J. Glieck, *Chaos: Making a New Science*, New York: Penguin. 1987.

[5] G. Monro and J. Pressing, *Sound Visualization Using Embedding: The Art and Science Of Auditory Correlation* Computer Music Journal, Vol. 2, No. 22, pp. 20-34. Summer 1998.

[6] F. Moon, *Chaotic Vibrations*, New York: Wiley. 1987.

[7] H. Olsen, *Music, Physics and Engineering*, New York: Dover. 1967.

[8] C. Roads, *The Computer Music Tutorial*, Cambridge: MIT Press. 1996.

[9] M. R. Schroeder, *Fractals, Chaos, Power Laws: Minutes from an infinite paradise*, New York: W.H.Freeman. 1991.

[10] B. Truax (ed.), *Handbook for Acoustic Ecology*, Vancouver: A.R.C. Publications. 1978.