X

AN ANALYTICAL APPROACH TO JOHN CHOWNING'S *PHONÉ*

Reiner Krämer, B.M.

Thesis Prepared for the Degree of

MASTER OF MUSIC

UNIVERSITY OF NORTH TEXAS

May 2010

APPROVED:

David B. Schwarz, Major Professor
Andrew May, Committee Member
Paul E. Dworak, Committee Member
Eileen M. Hayes, Chair of the Division of Music History, Theory, and Ethnomusicology
Graham H. Phipps, Director of Graduate Studies in the College of Music
James Scott, Dean of the College of Music
Michael Monticino, Dean of the Robert B. Toulouse School of Graduate Studies

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

INTRODUCTION

The Pythagoras Myth

> Pythagoras had long sought the rational criteria that determined musical consonances. One day, by divine guidance, he passed a smithy from which sounds of musical harmonies emerged. He approached the place with amazement, for pitches sounding consonant with each other seemed to come from the hammers. He examined the weights of the hammers and discovered that one weighed 12 pounds, a second 9 pounds, a third 8 pounds, and a fourth 6 pounds. The hammers of 12 pounds sounded the octave – that interval in which the two pitches were most identical. The hammers of 12 and 8 pounds, as well as those of 9 and 6 pounds, sounded the fifth – an interval which, next to the octave, was most beautiful. The hammers of 12 and 9 pounds, as well as those of 8 and 6 pounds, sounded the fourth – that interval which seemed to be the smallest consonance. In this manner Pythagoras discovered the ratios – the immutable essences – of musical harmonies: the octave lay in the ratio of 2:1; the fifth was determined by the ratio of 3:2; and the fourth was found in the ratio 4:3. Moreover, since the basic building block of music, the tone, was the difference between a fourth and a fifth, the ratio of that interval was the difference between 3:2 (or 12:8) and 4:3 (or 12:9), thus 9:8.
>
> Boethius, *De institutione musica*, liber I, Chapter 10[1]

---

[1] Calvin M. Bower, "The Transmission of Ancient Music Theory into the Middle Ages " in *The Cambridge History of Western Music Theory*, ed. Thomas Christensen (New York: Cambridge University Press, 2007), 142-143. Bower paraphrased the 10th chapter. The Latin version reads:

> Cum interea divino quodam motu praeteriens fabrorum officinas, pulsos maleos exaudivit, ex diversis sonis unam quodammodo concinentiam personare. Ita igitur ad id quod diu inquirebat attonitus, accessit ad opus: diuque considerans, arbitratus est diversitatem sonorum ferientium vires efficere. Atque ut id apertius colliqueret, mutarent inter se malleos imperavit. Sed sonorum proprietas non in hominum lacertis haerebat, sed mutatos malleos comitabatur. Ubi igitur id animadvertit, malleorum pondus examinat. Et cum quinque essent forte mallei, dupli reperti sunt pondere qui sibi secundum diapason consonantiam respondebant. Eumdem etiam qui duplus esset alio, sesquitertium alterius comprehendit, ad quem scilicet diatessaron sonabat. Ad alium vero quemdam, qui eidem diapente consonantia jungebatur, eumdem superioris duplum reperit esse sesquialterum. Duo vero hi, ad quos superior duplex sesquitertius et

A Challenge

New music presents new challenges to the field of music theory. The last 60 years, following the end of WWII, have seen the rise of electroacoustic music. The integration of computers within electroacoustic music began when Max Matthews was able to synthesize sounds on the computer in 1957 at Bell Telephone Laboratories Murray Hill, New Jersey.[2] Many bodies of music literature such as those composed in the Renaissance, the common practice period or the Second Viennese School, do have well-established analytical procedures. Electroacoustic music does not at present writing have well-established analytical procedures. It is difficult to establish any type of analytical language for new music. This makes it problematic to scope an analytical project in terms of what one might be looking for or what may be of interest for an analysis.

> sesquialter esse probatus est, ad se invicem sesquioctavam proportionem perpensi sunt custodire. Quintus vero est rejectus, qui cunctis erat inconsonans. Cum igitur ante Pythagoram consonantiae musicae, partim diapason, partim diapente, partim diatessaron, quae est consonantia minima, vocarentur primus Pythagoras hoc modo reperit, qua proportione sibimet haec sonorum chorda jungeretur. Et ut sit clarius quod dictum est, sint, verbi gratia, malleorum quatuor pondera, iquae subterscriptis numeris contineatur, 12, 9, 8, 6. Hi igitur mallei, qui 12 et 6 ponderibus vergebant, diapason in duplo concinentiam personabant. Malleus vero 12 ponderum ad malleum 9, et malleus 8 ponderum ad malleum 6 ponderum, secundum epitritam proportionem diatessaron consonantia jungebatur. Novem vero ponderum ad 6, et 12 ad 8 diapente consonantiam permiscebant. Novem vero ad 8, in sesquioctava proportione resonabant tonum.

Anicius Manlius Severinus Boethius, "De Institutione Musica, Liber 1", Indiana University http://www.chmtl.indiana.edu/tml/6th-8th/BOEDIM1_TEXT.html (accessed 03.08.2010).

[2] Joel Chadabe, *Electronic Sound: The Past and Promise of Electronic Music* (Upper Saddle River, New Jersy: Prentice-Hall, 1997), 108. Chadabe asserts that it was not until the rise of microprocessors and their widespread inexpensive use in non-mainframe computers that made computer music more accessible beginning in the late 1980s.

One of the problems music theorists encounter is what type of analytical method should be used. At the beginning of the 21^{st} century some new analytical methods have emerged that specifically deal with the analysis of electroacoustic music. These are presented in Thomas Licata's edited essays titled *Electroacoustic Music*.[3] Another book is Mary Simoni's edited collection of analytical essays titled *Analytical Methods of Electroacoustic Music*.[4] Leigh Landy's *Understanding the Art of Sound Organization* is one of the most recent sources of analytical methods.[5] Additional sources for analytical methods can also be found in journals such as the *Computer Music Journal*, *Organised Sound*, *Leonardo*, *Perspectives of New Music* and *Journal of Music Theory*.

Another group of sources that aid in analysis are so-called technique or tutorial books. The technique books describe how to program and manipulate computer-generated sounds in respect to techniques that previously have been used by researchers, programmers, and composers of computer music. The composer Charles Dodge and the researcher Thomas A. Jerse published one of the earliest technique books titled *Computer Music* in 1985. An upgraded second edition was published in 1997.[6] In 1996, the composer Curtis Roads published *The Computer Music Tutorial*.[7] He also published a book titled *Microsound* in 2001 and has another tutorial book titled *Composing Electronic Music* that is

---

[3] Thomas Licata, ed. *Electroacoustic Music* (Westport, Connecticut: Greenwood Press, 2002).
[4] Mary Simoni, ed. *Analytical Methods of Electroacoustic Music* (New York: Routledge, 2006).
[5] Leigh Landy, *Understanding the Art of Sound Organization* (Cambridge: MIT Press, 2007).
[6] Charles and Thomas A. Jerse Dodge, *Computer Music*, 2nd ed. (New York: Schirmer Books, 1997).
[7] Curtis Roads, *The Computer Music Tutorial* (Cambridge, Massachusetts: The MIT Press, 1996).

expected to be released in 2010.[8] The mathematician Miller Puckette published *The Theory and Technique of Electronic Music* in 2007.[9] All books mentioned above feature exhaustive computer music terminology indices. However, the most current and extensive source for computer music terminology and techniques is the *Global JOS Index* website assembled by Julius O. Smith III.[10] This index website is constantly updated and adds new computer music terminology, techniques and related terminologies (acoustics, mathematics) as they emerge.[11] Furthermore, John Chowning explains that techniques listed in the *JOS index* should be linked to MaxMSP/Pd patches.[12]

Computer language manuals also serve as guides for the music theorist. Computer music exists for more than 50 years now (as long as one of the oldest programming languages named LISP developed in 1958), which caused many common programming practices to grow together into code libraries that are used to program music synthesis on the computer. Max Mathews developed a programming environment called MUSIC in 1957, which had grown into MUSIC

---

[8] Curtis Roads, *Microsound* (Cambridge: MIT Press, 2001).
[9] Miller Puckette, *The Theory and Technique of Electronic Music* (Hackensack, NJ: World Scientific, 2007).
[10] Julius O. Smith III, "Global Jos Index", Center for Computer Research in Music and Acoustics, Stanford University https://ccrma.stanford.edu/~jos/GlobalJOSIndex.html (accessed 03.07.2010).
[11] John Chowning has “long been impressed with Julius Smith’s Global JOS Index and its ever-evolving context on the CCRMA server.” John M. Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future," in *Computer Music Modeling and Retrieval. Sense of Sounds: 4th International Symposium, Cmmr 2007, Copenhagen, Denmark, August 27-31, 2007. Revised Papers*(Springer-Verlag, 2008), 6.
[12] Ibid., 9.

11 by the early 1980s.[13] Chowning used various incarnations of MUSIC and many of his algorithms were written in this environment.[14]

With the emergence of the C programming language in the 1980s, Richard Moore developed a C port for Music 11 in 1985 at UCSD's CARL called CMUSIC, which was followed by Barry Vercoe's development of CSound in 1986 at MIT.[15] Both of these environments contain ports of Chowning's music synthesis algorithms. Perry R. Cook (Princeton University) and Gary P. Scavone (McGill University) have ported some of Chowning's code or variations thereof to *The Synthesis ToolKit in C*++, or STK for short, which sits on Stanford's CCRMA server.[16] Synthesizer manuals are also useful, especially when authored by the composer.[17] Additionally the aforementioned journals also contain information on compositional and generative or manipulative sound techniques as they are emerging.[18]

Consulting computer music technique books and language manuals are good points of departure to investigate musical ideas and practices of the genre. However, when analyzing music it is equally important to experiment with these

---

[13] Peter Manning, *Electronic and Computer Music* (New York: Oxford University Press, 2004), 187-189.
[14] Ibid., 189.
[15] UCSD--University of California, San Diego. CARL--Computer Audio Research Laboratory. "Port" refers to a translation from one programming language to another. CSound is still in use in the present. Ibid.
[16] *Center for Computer Research in Music and Acoustics*. John Chowning was one of the directors and founders of CCRMA. *ChucK*, a synthesis programming language by Ge Wang (Stanford University), utilizes the STK.
[17] David Bristow and John Chowning, *Fm Theory and Applications* (Tokyo: Yamaha Music Foundation, 1986).
[18] Puckette.

musical ideas and practices. In non-electroacoustic music, it makes sense to use a printed score, the piano, one's voice and staff paper.[19] Since computer music is written and performed, alone or interactively with other musicians, on or with the computer it makes sense to use a computer as an addition to one's lab tools. The computer itself needs to be equipped with programming tools. There are countless computer music environments to create computerized audio synthesis.

Several of them (CSound, the STK, ChucK) have already been mentioned. Pd, also known as PureData, is a visual programming environment and framework.[20] Pd in itself can be seen as a musical instrument. Like many musical instruments it has the ability to change tone color, timbre, etc. but can do even more than a traditional musical instrument since it can be built in such a fashion that it can provide a model of musical intelligence. Pd can also act as a tape recorder, a digital reverberator, a synthesizer, or any of a variety of well-known sound appliances. Within this environment, one can create patches, small

---

[19] This practice in itself is a form of modeling.

[20] Other visual programming frameworks for computer music also exist. They are mainly jMax, developed at IRCAM (Institut de Recherche et Coordination Acoustique Musique) and Max/MSP commercially developed by Cycling 74. These environments grew out of software developed at IRCAM in the 1980s-90s. Miller Puckette was an active developer for work done on these environments in the 1980s-90s. He is also the main developer of Pd. Preference of Pd is given here. For one, Pd is freely available to download from Miller Puckette's web site (http://crca.ucsd.edu/~msp/software.html). Second, Miller Puckette illustrates his examples in his technique book with Pd. Third; Pd has a large active open source software development community behind it, which enables the environment to run smoothly on any computer system. Because this software is non-commercial, it will not rise and fall with a life cycle of a private company.

visual computer programs that represent logic and calculations that can affect or produce sound.[21]

Patch environments have existed now for twenty years and are rooted within electronic and computer music itself. With Pd, it is possible to re-create any aspect of electronic or computer music from the turn of the century up to the present. Early electronic music tools and instruments can be recreated, from simple oscillators to complex synthesizers and signal processors. Pd represents a medium that can contain all aspects of electro-acoustic music and can recreate any electroacoustic experiment and experience from the past.[22]

In addition to Pd, the music theorist's lab needs to include software tools like a sound editor, a spectral visualizer and a spectral editor.[23] The sound editor used in this project needed to have the ability to create markers in time alongside a sound file and needed to have the ability to export smaller sound samples

---

[21] Patches created in Pd are really shorthand versions of more complex programming procedures in the C language. A composer, with some practice, can easily sketch out musical ideas intuitively, rather than spending a lot of time learning complex programming procedures. The patches can be seen as prototypes, from which programs can be developed in other programming languages as well.

[22] The history of the Max family (which contains jMax, Max/MSP and Pd) starts with the Patcher editor written by Miller Puckette. This editor was specifically written for Philippe Manoury's composition Pluton. The main focus of creating this editor was to create a GUI (Graphic User Interface) that would be easy to use for non-programmers, mainly composers. Originally used to control IRCAM's unique 4X synthesizer, by 1989 Max had evolved into a hardware/software system using NeXT computers and ISPW signal processing cards. By 1995, the speed of processors on consumer computers was sufficient that no specialized processing hardware was required; in the succeeding years, Miller Puckette created Pd and David Zicarelli released the commercial software Max/MSP; IRCAM used the newly developed Java environment to create J-Max, the third member of the Max family. More about the history of Max can be read in Miller Puckette, "Max at Seventeen," *Computer Music Journal* 26, no. 4 (2002).

[23] Generally, a notation software tool should also be in the music theorist's toolbox, but is not as relevant here. All tools mentioned, with the exception of the notation software, can also be assembled within Pd.

encapsulated by these markers. *Audacity*, which is a freely available open-source audio editor, has these attributes.[24] The spectral visualizer needed to have the ability to plot different spectral analyses of previously exported sound samples. One such tool is *Sonic Visualizer,* which is developed at the Centre for Digital Music, Queen Mary, University of London.[25] Michael Klingbeil created a *Sinusoidal Partial Editing Analysis and Resynthesis* tool known by its acronym *SPEAR*.[26] *SPEAR* was used in this project because it tracks a very detailed spectral envelope that can be manipulated (for example: a given sound sample can be slowed down, sped up or its intensity envelope can be increased or decreased and individual partials can be played back).

These software tools are of great aid to an analyst, but it is also important to know histories of electroacoustic music. There are many choices. Joel Chadabe's *Electric Sound* is a history of electroacoustic music narrated through the eyes of the creators of electroacoustic music.[27] Another historical source is Peter Manning's *Electronic and Computer Music*.[28] A more general approach not specifically dealing with electroacoustic music can be found in Robert P. Morgan's *Twentieth-Century Music: A History of Musical Style in Modern Europe*

---

[24] *Audacity* is available at http://audacity.sourceforge.net/.
[25] *Sonic Visualizer* is available at http://www.sonicvisualiser.org/.
[26] *SPEAR* is available at http://www.klingbeil.com/spear/ and is based on Michael Kateley Klingbeil, "Spectral Analysis, Editing, and Resynthesis: Methods and Applications" (D.M.A. diss., Columbia University, 2009).
[27] Chadabe.
[28] Manning.

and America.[29] A specific historiography is Évelyne Gayou's edited collection of *John Chowning* essays.[30] Additional sources for historical perspectives can be found in yearly reports of state funded institutions.[31]

Collected essays on the general subject of computer music by its practitioners provide insights to aesthetical ideas of composers. One of these is *Current Directions in Computer Music Research* edited by Max Matthews and John Pierce.[32] Another one is Perry Cooke's *Music Cognition and Computerized Sound*.[33] Furthermore, conference papers and proceedings offer insights into more sets of principles developed by computer music composers. John Chowning presented one of these proceedings papers, titled *Digital Sound Synthesis, Acoustics and Perception: A Rich Intersection*, in December 2000.[34] He presented another one, titled *Fifty Years of Computer Music: Ideas of the Past Speak to a Future – Immersed in Rich Detail*, in August 2007.[35]

Seminars, program notes and interviews supply further information on aesthetic approaches of computer music practitioners. Such a collection of

---

[29] Robert P. Morgan, *Twentieth-Century Music: A History of Musical Style in Modern Europe and America* (New York: W. W. Norton, 1991).
[30] Évelyne Gayou, ed. *John Chowning*, 12 vols., Portraits Polychrome, vol. 7 (Paris: Institut national de l'audiovisuel, 2007).
[31] *Hamburger Jahrbuch Für Musikwissenschaft*. vol. 11 (New York: P. Lang, 1991). A *Jahrbuch*, literally meaning 'year book', is usually a yearly report of activities and research published by a publically funded German university or research institution.
[32] Max V. Mathews and John R. Pierce, eds., *Current Directions in Computer Music Research*, System Development Foundation Benchmark Series (Cambridge, Massachusetts: The MIT Press, 1989).
[33] Perry R. Cook, ed. *Music, Cognition, and Computerized Sound : An Introduction to Psychoacoustics* (Cambridge, Mass.: MITPress, 1999).
[34] John M. Chowning, "Digital Sound Synthesis, Acoustics and Perception: A Rich Intersection," in *COST G-6 Conference on Digital Audio Effects* (Verona, Italy: 2000).
[35] Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future."

papers and program notes was compiled after seminars and concerts organized by IRCAM and *l'Ensemble InterContemporain* that were held at IRCAM in Paris between the 17th and 21st of February 1981.[36] An additional one is *Composers and the Computer*, which features an interview with John Chowning conducted in April 1982.[37] Moreover, the journals previously mentioned also provide additional historical information.

These analytical approaches, technique books, theory books, programming manuals, articles, histories, essays, conference proceedings and papers, program and seminar notes, and yearly reports present much material as a point of departure for analysis. All of these resources are relevant to the understanding of how and why computer music compositions work.

---

[36] Marc Battier and others, eds., *Le Compositeur Et L'ordinateur* (Paris: IRCAM en association avec l'Ensemble InterContemporain, 1982). The concert was held around the time Chowning had finished composing *Phoné*. L'Ensemble InterContemporain was formed in 1976 by Pierre Boulez and the French Ministry of Culture. The ensemble is based at the *Cité de la Musique* in Paris since 1995 and is currently active and continues to be supported by the Minisitry of Culture and Communication. More information is available at http://www.ensembleinter.com/.

[37] Curtis Roads, ed. *Composers and the Computer* (Los Altos, Calif.: W. Kaufmann, 1985). This interview was help about one year after Chowning had finished composing *Phoné* and upon his return to Stanford. Interestingly enough it has the same title in English, but only contains interviews rather than essays.

The Composition

From the early days of electroacoustic music until the present, interaction of electronically generated sounds with human vocal sounds or altered human vocal sounds has been a favorite among composers within this genre of music. Compositions that exemplify this practice are *Gesang der Jünglinge* (Stockhausen, 1955-6), *Epitaph für Aikichi Kuboyama* (Eimert, 1957), *Thema-Omaggio à Joyce* (Berio, 1958), *Nouvelles Aventures* (Ligeti, 1962-65), *Speech Songs* (Dodge, 1976), *Soft Morning, City!* (Machover, 1980), *Six Fantasies on a Poem by Thomas Campion* (Lansky, 1978-79), *Mortuos Plango, Vivos Voco* (Harvey, 1980) and *Phoné* by John Chowning in 1980/81.[38]

Phoné was realized at Stanford University in 1980/81, after Chowning had returned from IRCAM.[39] The composition "was premiered at IRCAM in Paris in February of 1981."[40] In *Phoné*, Chowning utilized an expansion of a sound synthesis technique called frequency modulation synthesis (FM synthesis – Figure 1).[41]

---

[38] David Evan Jones, "Compositional Control of Phonetic/Nonphonetic Perception," *Perspectives of New Music* 25, no. 1/2 (1987): 139.
[39] John Chowning, "DVD Program Notes," *Computer Music Journal* 32, no. 4 (2008): 109.
[40] Ibid.
[41] Julius Smith's definition of FM synthesis:

> A general formula for frequency modulation of one sinusoid by another can be written as
>
> $$x(t) = A_c \cos[\omega_c t + \phi_c + A_m \sin(\omega_m t + \phi_m)],$$
>
> where the parameters (A_c, ω_c, ϕ_c) describe the carrier sinusoid, while the parameters (A_m, ω_m, ϕ_m) specify the *modulator* sinusoid … modulation of phase implies a modulation of frequency, and vice versa, since the instantaneous

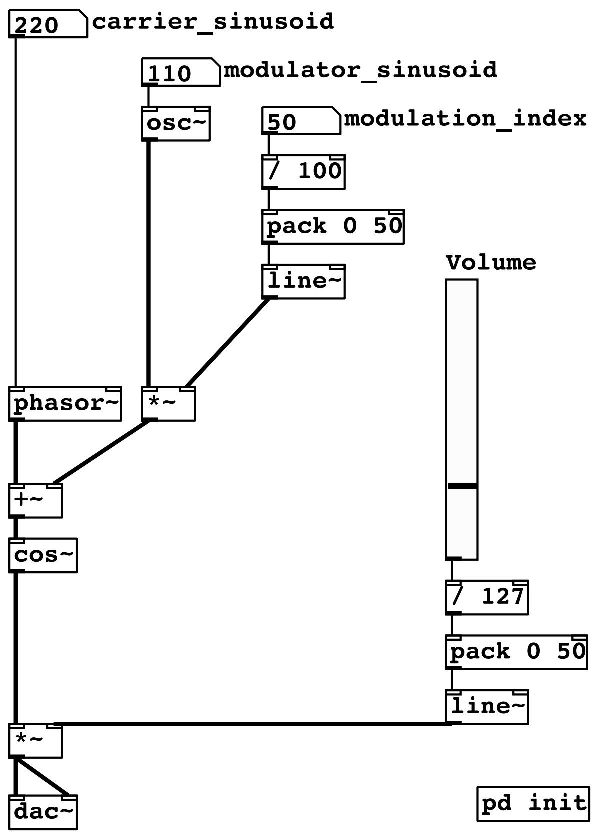

Figure 1: FM synthesis as described in the JOS index.

---

frequency is always defined as the time-derivative of the instantaneous phaseJulius O. Smith III, "Sinusoidal Frequency Modulation (Fm)", Center for Computer Research in Music and Acoustics, Stanford University https://ccrma.stanford.edu/~jos/mdft/Sinusoidal_Frequency_Modulation_FM.html (accessed 03.07.2010).

Figure 1 demonstrates how the mathematical formula from above can be expressed in a working *Pd* patch. The carrier sinusoid can be changed to any specified value in Hz. The modulator sinusoid can be changed to any specified value in Hz. Two parameters were added. One is used to control the volume of the sound output of the digital audio converter (`dac~`) – i.e. a multiplier (`*~`) is controlled by the output of a volume slider that is ramped over 50 milliseconds (`pack 0 50`, `line~`) in order to avoid "clicking" sounds while changing the volume. The second is the modulation index that controls how much of the modulator sinusoid will be added to the carrier sinusoid. Another ramp is used (`pack 0 50`, `line~`) to avoid further "clicking" sounds with the modulation index. The contents and a short description of the `pd init` object are in APPENDIX B – PD PATCHES, Figure 36. Text that is part of a given text patch will use `courier` font.

Chowning was the first composer and researcher to utilize linear FM with a dynamic index of modulation in 1967 to produce computer generated sounds. He had previously used this technique in his compositions: *Sabilithe* (1966, revised 1971), *Turenas* (1972) and *Stria* (1977). Each composition added another technique to the canon of computer music. *Sabilithe* represented an experimentation with the first digitally FM synthesized sounds. *Turenas* dealt with the spatial perception of sound. *Stria* employed an alternate tuning system based on the golden mean and *Phoné* utilized the FM synthesis of a human-like voices. The FM synthesis algorithm that Chowning used in *Phoné* was developed at IRCAM in Paris, 1979.[42] Chowning's recent composition *Voices* (2007) exploited FM synthesis and follows along the tradition of compositions for voice and computer. *Voices* was realized within the Max/MSP environment.[43]

Chowning's work of creating digital FM synthesis left a large impact on music. His algorithms were commercially licensed through Stanford University to the Yamaha Corporation, which used Chowning's digital creations in its line of Yamaha DX synthesizers throughout the 1980's.[44] In effect, digital sound synthesis was not confined to a specialized lab and became inexpensive. Chowning states on the importance of FM synthesis in music:

> FM synthesis led to new thoughts about sound spectra, tuning and pitch. Detached from their traditional association with the timbre of acoustic instruments, spectra become structured and associated

---

[42] Chowning, "DVD Program Notes," 109.
[43] Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future," 7.
[44] Richard F. Moore, "Dreams of Computer Music: Then and Now," *Computer Music Journal* 20, no. 1 (1996): 35-36.

> with pitch in ways that are unique in the medium of computer music.[45]

François Rose saw Jean-Claude Risset's work on spectral envelopes and Chowning's work on FM synthesis as a direct prologue to the development and understanding of spectral music.[46] Viviana Moscovich seconded this argumentation.[47] Joshua Fineberg followed along the same thread.[48]

The historical lineage, the technique (FM synthesis, reverb, etc.), the impact on the continually expansion of musical timbre, and the influence on other compositional techniques (spectral music) placed *Phoné* into context with acoustic music, electroacoustic music and computer music.[49]

---

[45] Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future," 1.
[46] François Rose, "Introduction to the Pitch Organization of French Spectral Music," *Perspectives of New Music* 34, no. 2 (1996).
[47] Viviana Moscovich, "French Spectral Music: An Introduction," *Tempo* 200 (1997).
[48] Joshua Fineberg, "Guide to the Basic Concepts and Techniques of Spectral Music," *Contemporary Music Review* 19, no. 2 (2000).
[49] "Acoustic music" as in music to be performed by 'traditional' wind, string and/or percussion instruments.

The Object

> Whatever happens in a piece of music is the endless reshaping of the basic shape…. There is nothing in a piece of music but what comes from the theme, springs from it and can be traced back to it.[50]
>
> Arnold Schoenberg

The “basic shape” is known as *Grundgestalt* and according to Michael J. Schiano’s definition in the New Grove “may be a fragment of the musical surface that subsequently undergoes repetition, variation, development and ‘liquidation’ as the piece unfolds.”[51] Schiano further explains that “for Schoenberg the *Grundgestalt* was…a construct transcending stylistic distinction.”[52]

Electroacoustic music and computer music composers have adapted a similar concept. Jerome Kohl testified, “*Telemusik* must…be described…as a moment form, made up primarily of group and statistical (gestalt [sic] and textual) elements….”[53] Agostino Di Scipio showed, “[Jean-Claude] Risset projects the gestalt [sic] properties of the pitch contour into the realm of timbre.”[54] Landy quoted from Pierre Schaeffer’s *Traité des objet musicaux*: “The sound

---

[50] Arnold Schoenberg, "Linear Counterpoint," in *Style and Idea*, ed. Leonard Stein(Berkeley: University of California Press, 1984), 290.

[51] Michael J. Schiano, "Grundgestalt", Grove Music Online. Oxford Music Online. http://www.oxfordmusiconline.com/subscriber/article/grove/music/11868 (accessed 10.06.2009). Schoenberg described liquidation in *The Musical Idea and the Logic, Technique, and Art of its presentation* as the process of “gradually eliminating characteristic features, until only uncharacteristic ones, remain, which no longer demand a continuation.” Schoenberg further states, “Often only residues remain, which have little in common with the basic motive.” Arnold Schoenberg, *The Musical Idea and the Logic, Technique and Art of Its Presentation*, ed. Patricia Carpenter and Severine Neff, trans., Patricia Carpenter and Severine Neff (Bloomington, Ind.: Indiana University Press, 2006), 264.

[52] Schiano.

[53] Jerome Kohl, "Stockhausen: Telemusik (1966)," in *Electroacoustic Music*, ed. Thomas Licata(Westport, Connecticut: Greenwood Press, 2002), 113.

[54] Agostino Di Scipio, "Jean-Claude Risset: Contours (1982)," in *Electroacoustic Music*, ed. Thomas Licata(Westport, Connecticut: Greenwood Press, 2002), 177.

object…can be compared to a 'gestalt' [sic] in the psychology of form.'"[55] Trevor Wishart explained that in his composition *Red Bird*, "by a gradual shaping of the envelope and the addition of reverberation the character of the gestalt is completely altered."[56] Chowning referred to the Gestalt 'law of common fate' while speaking of Risset's work on trumpet tones in the late 1960s and early 1970s.[57] He explains:

> The evolution of the harmonic amplitudes of the spectrum during the arrack portion is rapid, complicated, but patterned. As the overall amplitude of the tone the greater the relative contribution of the higher harmonics….The pattern, however, is not discernable by the "ear" as it is heard as a totality according to the Gestalt "law of common fate" where components moving in the same direction are perceived as a group.[58]

Leigh Landy put *Phoné* directly into a psychological context that can be related to the gestalt principle:

> There are pieces that play with what the Germans call *Schein und Sein*, semblance and reality. John Chowning's Phoné exemplifies this tendency in a work where the listener is left guessing which sounds are vocal and which synthesized. He seems to be able to act as a magician manipulating the listener's perception.[59]

Schein und Sein is the guiding force of the basic shape or object defined as X in John Chowning's composition *Phoné*.[60] X occurs multiple times in *Phoné*.[61]

---

[55] Pierre Schaeffer, *Traité Des Objets Musicaux: Essai Interdisciplines*, 2nd ed. (Paris: Seuil, 1977). Quoted in Landy, 80.
[56] Trevor Wishart, *On Sonic Art*, ed. Simon Emmerson (Amsterdam: Harwood Academic Publishers, 1996), 174.
[57] Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future," 2.
[58] Chowning, "Digital Sound Synthesis, Acoustics and Perception: A Rich Intersection," 2.
[59] Landy, 32.
[60] X is given in order to avoid epistemological and ontological problems that may arise with the reuse of Schoenberg's term of *Grundgestalt* or Pierre Schaeffer's concept of *objet musicaux*. X refers to a set of control parameters.

In the following study, I shall examine 4 occurrences of X in detail in John Chowning's *Phoné*. In the following chapter called *Perspectives,* I shall discuss the historical background of the composition and under what circumstances it came to be. Further, I shall propose what constitutes X in *Phoné*. Then, I shall briefly discuss the form of the piece and how instances of X fit into a larger picture. In the following chapter called *X-Events* I shall show how each of these incidences map onto X by examining pitch, timbre, harmonic and rhythmical material used by Chowning. I shall pay particular attention to timbre and its developing spectral envelopes through the use of Fast Fourier Transform analysis (FFT analysis). I shall then study the results of the FFT analyses and shall show how these analyses are visualized as spectrograms in *Sonic Visualizer* and *SPEAR*. Next, I shall evaluate these results and shall propose synthesis techniques as music theoretical entities.[62] Subsequently, I shall provide different prototypes in *Pd* that help to *auralize* these concepts.[63]

---

[61] A complete list of proposed X material is listed Table 10: X Occurrences.
[62] A "music theory entity" is anything that is used to describe any type of musical technique with or without function, e.g.: A V^7 Chord, a voice exchange, a cross-relationship, a chord substitution, a cadence, etc.
[63] *Auralize* is used in the sense of *visualize*. The verb was described by Steve Larsen, "The Problem of Prolongation In "Tonal" Music: Terminology, Perception, and Expressive Meaning," *Journal of Music Theory* 41, no. 1 (1997): 104, 132. Furthermore, the term was used by Brian Fennelly, "A Descriptive Language for the Analysis of Electronic Music," *Perspectives of New Music* 6, no. 1 (1967): 85, 92.

PERSPECTIVES

X

John Chowning had started to work with one of his students, Michael McNabb, on the synthesis of vocal sounds at Stanford in 1978.[64] Chowning was invited to work at IRCAM in Paris from late 1979 to early 1980, where he continued to work on vocal synthesis utilizing his frequency modulation technique. At IRCAM Chowning was influenced by Johann Sundberg's studies of vocal formants, which quickly led to the development of algorithms for the synthesis of vocal sounds.[65] Sundberg's research work led to the development of CHANT, a project headed by Xavier Rodet, who stated that "CHANT was a physical model," in which sounds could be fine-tuned.[66] CHANT was an additive synthesis program and even though Chowning was the discoverer of FM synthesis, he did prefer to utilize additive synthesis first in his research of vocal synthesis.[67] Sundberg's studies at IRCAM were directly reflected in 2 of

---

[64] Bruno Bossis, "Phoné," in *John Chowning* , ed. Évelyn Gayou, 103-105 (Paris: Institut national de l'audiovisuel, 2007) 103.
[65] Bossis, 103; John M. Chowning, "Frequency Modulation of the Singing Voice," in *Current Directions in Computer Music Research* , ed. Max V. Mathews and John R. Pierce, 57-64 (Cambridge, Massachusetts: MIT Press, 1989).
[66] Joel Chadabe, *Electric Sound* (Upper Saddle River, New Jersey: Prentice Hall, 1997), 126.
[67] Bossis, 104.

Chowning's 7 observations, about soprano tones.[68] He credited Sundberg's article *Synthesis of Singing* in his first and third observation.[69]

One of Chowning's goals was to synthesize a vocal sound as "real" as possible and carefully observed nuances of the human voice, such as the quantity of random variation in the vibrato of portamento during the attack phase, the length, or the decay and sustain phase, and the release phase of a vocal sound.[70] While Chowning was working at IRCAM he noticed the perceptual ambiguities between his instrumental and vocal timbres that used periodic or random vibrato. One of the instrumental sounds Chowning used was a metallic bell-type sound. Chowning was interested in how this sound could be transformed into any other timbre.[71]

Chowning was not a singular user of sound transformations in France. Risset was interested in sound transformations as well. He stated, "Since 1979 I have done further experimentation…particularly on processes for sonic transformation and development."[72] Risset further noted that he used this research for his compositions *Contours*, *Profils* and *Aventure de lignes*. Jonathan Harvey worked on transforming vocal sounds at IRCAM in 1980. Harvey brought sampled sounds of a boy's voice and completed FFT analyses of these samples.

---

[68] John M. Chowning, "Frequency Modulation Synthesis of the Singing Voice," in *Current Directions in Computer Music*, ed. Max V. Mathews and John R. Pierce(Cambridge: MIT Press, 1989), 58-59.
[69] Johann Sundberg, "Synthesis of Singing," *Swedish Journal of Musicology* 60, no. 1 (1978).
[70] Roads, ed. *Composers and the Computer*, 22-23.
[71] Ibid., 23.
[72] Jean Claude Risset, "Computer Music Experiments 1964 - ... [sic]," *Computer Music Journal* 9, no. 1 (1985): 17.

He programmed CHANT to simulate the results of the FFT analyses. Harvey then proceeded to apply transformations that were “applied to the spectra of the boy’s vowels, which could be made into pitch and amplitude glissandi to the nearest bell equivalents in a bell spectrum.”[73]

Phoné is Chowning’s playground for timbral transformations. The title of the composition *Phoné*, old Greek for “voice”, reveals the nature of the timbral transformation. The bell sound will be transformed mainly to a human-like vocal timbre, with the aid of FM synthesis.[74] The timbral transformations are based on X.

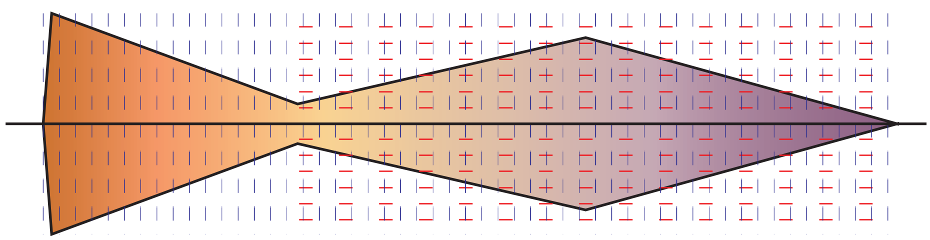

Figure 2: X in *Phoné*.

A visual representation of *Phoné*’s X (Figure 2) shows a metallic bell sound with a sharp attack followed by rapid decay. These initial sections will be referred to as the *head* sections. The metallic sound can evolve into several types of FM synthesis configurations. The most striking transformations are the FM vocal synthesis configurations that include the addition or simulation of formants. These transformations will be referred to as *tail* sections. The x-axis is

[73] Jonathan Harvey, "'Mortuos Plango, Vivos, Voco': A Realization at Ircam," *Computer Music Journal* 5, no. 4 (1981): 24.
[74] Transformations to other sounds also occur, but really spring to life from the same seed.

time and the y-axis is the intensity level. These FM constructions, being either metallic or vocal-like are represented here with dashed red lines. The lines are dashed to show that the sounds may begin and end at different event points. The horizontal spacing between the red dashed lines is arbitrary and not scaled to any actual pitch frequencies. The spacing should be thought of as being a placeholder; meaning that any pitch frequency can be inserted. The evolutions of the vocal-like FM sounds are dovetailed into the decay of the bell sounds. Next, the vocal-like FM sounds slowly swell, come to a plateau (sustain) and are followed subsequently by a slow decay.

The intensity envelope is also a representation of the timbral envelope. From this perspective, the x-axis is frequency in Hz and the y-axis remains to be the intensity level. The dashed purple lines show frequencies as they could occur in a spectral envelope. Again, the lines are dashed to demonstrate that these pitch levels in the spectral envelope have liquid and different intensities depending on the variation of X. The color shading and its transformation within the shape show the timbral transformation that occurs with each instance of the fantasy based upon X.[75] In order to map *Phoné*'s event terrain and place instances of X into a larger context an overview of its formal sections becomes necessary.

---

[75] Curtis Roads interviewed John Chowning on the afternoon of April 29th, 1982 at CCRMA (Center for Computer Research in Music and Acoustics at Stanford University, CA). Roads asked Chowning whether *Phoné* was "rigorously organized" or an "improvisation." Chowning's reply was that is was "right in the middle" and that "there was more fantasy in its composition." Roads, ed. *Composers and the Computer*, 23.

Mapping

Phoné has 4 major sections.[76] Their respective labels are: A, B, C and D. The A section assumes an expository role in the composition. A begins very subdued and increases in intensity to a climax. The first section (A-1) presents the musical problem and begins with a collection of rapidly increasing bell sound occurrences that transform into FM synthesized sounds. The transformations in A-1 seem not to evolve into clearly distinguishable vocal sounds. The A-1 section lasts for 42.72 seconds. The second section (A-2), which can be subdivided further into two sections, ends at 5:03 minutes. This section is a vast experimentation of how to synthesize vocal sounds and their formants. Many of the occurrences of X seem to shimmer in an out of metallic sound timbres. The contour of the entire A reflects the shape of X.

The B section follows that climax and functions as a bridge to the C section. B begins very similar to how the A section began, with a rapid bell cluster that is transformed into a complex synthesized sound. X assumes a more polyphonic role. However, after a few false beginnings, which are reminiscent of the opening gesture, Chowning's transformations of X become increasingly more

[76] The precise spelling varies within different sources. Computer Music Journal and the Wergo record label have used the former, whereas the latter is a transliteration of the old Greek *Φωνή* used by Curtis Roads. The transliteration of *Phoné* will be used here. These four sections have been previously described by Bossis, however, closer approximations of delimiters are presented here. Bruno Bossis, *Phoné*, Vol. VII, in *John Chowning Portraits polychromes*, ed. Évelyne Gayou, 112 (Paris: Institut national de l'audiovisuel, 2007): 103-105. The Institut national de l'audiovisuel produced a website about John Chowning and analyses of 3 of his compositions. The formal sections by Antonio Salluce are congruent with the ones described by Bruno Bossis. Antonio Salluce, "John Chowning", Institut national de l'audiovisuel http://www.ina-entreprise.com/sites/ina/medias/upload/grm/portraits-polychromes/extraits/chowning/index.html (accessed 03.07.2010).

distinguishable. At one point, he uses X with canonic imitations that are split among the audio channels.

The end of C begins with a climax that ebbs off in intensity. X transcends actual human physicality and becomes part of the physicality of the computer, which is a transformation itself. Chowning explains:

> The interpolation between ‘real’ timbres into registers that could not possibly exist in the ‘real’ world, such as the basso “profundissimo,” and the microstructural control of sound having to do with perceptual fusion and segregation of spectral components, are important points in this composition.[77]

The composer underlines one of the main aesthetic points of why computer music is powerful: to synthesize sounds that sound “real” or seem to sound real but are not. C contains what Chowning and others have referred to as the *basso profundissimo* vocal timbres.

D closes the composition and functions as a coda. The composer uses his X almost exclusively in this section. Multiple occurrences and variations of X happen all the way to the end and ebb off in intensity levels. The *coda* (D) winds the velocity of the composition down. The coda functions as playground of what Schoenberg calls ‘liquidation’ which further underlines the existence of X.

These sections serve as coordinates for events of interest in Phoné and are summarized in Table 1:

Section	Start Time	End Time	Duration
A	00:00	05:06	05:06

[77] Chowning, "DVD Program Notes," 109.

Section	Start Time	End Time	Duration
B	05:06	08:03	02:57
C	08:03	11:18	03:14
D (Coda)	11:18	13:00	01:42

Table 1: *Phoné*'s formal sections in minutes and seconds.

Figure 3 shows a spectrogram of the entire composition. The x-axis is the duration of the piece in time (the minutes are indicated on top of the spectrogram). The y-axis represents the frequency range in Hz, with peak regions at around 6000 Hz. The darker colors indicate higher intensity levels or dynamic levels in dB.

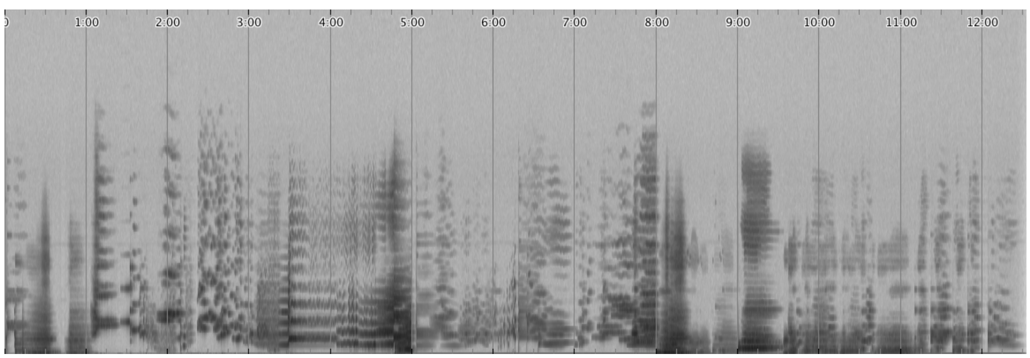

Figure 3: *Phoné* Spectrum.

Figure 4 shows the waveform in connection with pitch levels.[78] The y-axis here shows the intensity levels.[79] The background shows the pitch level on the y-axis and the shading illustrates intensity levels. When the shade of the pitch levels is darker, the intensity level of the pitch material is stronger. The second

[78] A picture of a waveform similarly to Figure 3 can also be generated in Pd. It is most helpful when dealing with small samples. APPENDIX B - *PD* PATCHES, Figure 37 shows how to build an audio player that can display the waveform.
[79] The terms "Intensity level" or "dynamic level" are used here interchangeably.

figure also illustrates the formal section of the composition in association with Table 1. Higher intensity levels occur either at the beginning or the end of the formal sections.

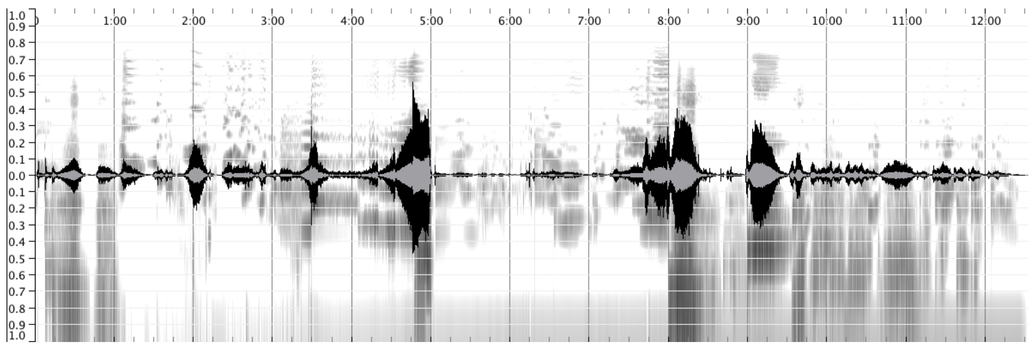

Figure 4: *Phoné* spectrum with waveform.

Unity is achieved in these formal sections, because each of the sections makes use of at least one incarnation (or fantasy) of X. Variety is achieved because each incarnation of X is different. It is now possible to properly investigate X occurrences of interest (micro level) within the four-part form (macro level).

X-EVENTS

Preliminary Gesture

Phoné begins with a series of 13 frenzied accelerating pitch events (13er group) that are followed by a 14^{th} sustained pitch event, resembling a metallic sound.[80] The metallic sound is the subject of transformation and makes up the head of following Xs. It is 6.5 seconds long and is the first gesture of the composition. This gesture is followed by X-1.

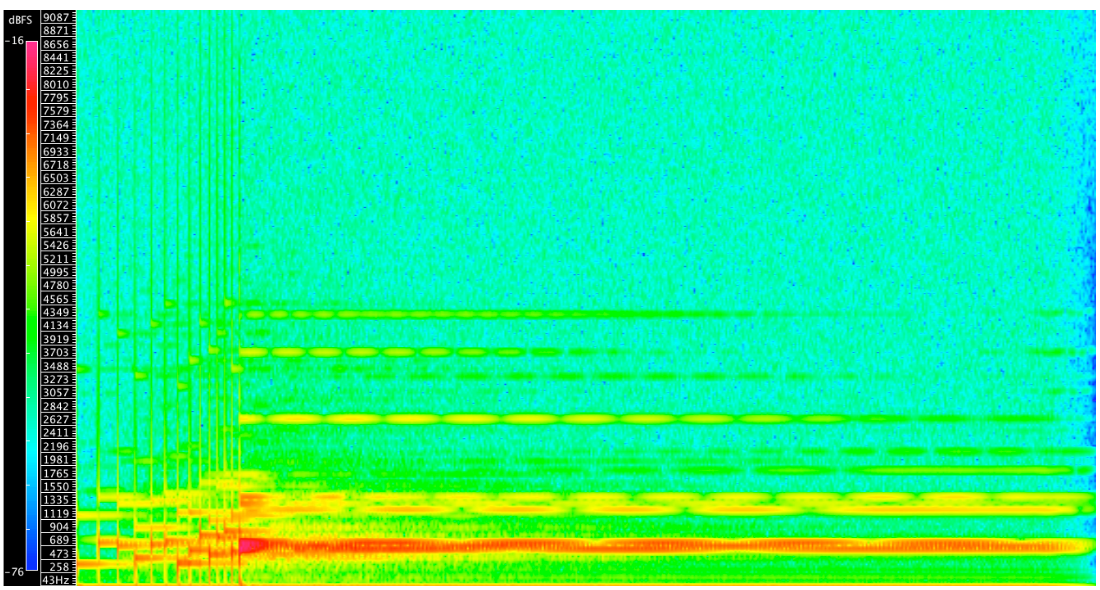

Figure 5: Spectrogram of the first 6 seconds of the composition.

[80] *Pitch events* can consist of a perceived single note and its corresponding harmonic spectrum, or a set of clustered notes that occur simultaneously and their corresponding harmonic spectra. “13er group” refers to the first 13 notes of the composition. The following spectrogram was realized using *Sonic Visualizer*.

Figure 5 shows a spectrogram of this preliminary gesture.[81] Notice rising vertical lines, illustrating the sharp attacks of each of the 14 pitch events and how these attacks affect the surrounding frequencies. The y-axis shows a frequency scale in Hertz (Hz). The x-axis is phenomenal time. The width of Figure 5 spans over a duration of ca. 6.5 seconds. The color shading indicates the intensity levels of the corresponding pitch events according to the scale on the left side of the figure. The figure shows where the pitch events appear and reveals the nature of these pitch events.

In a preliminary investigation, one observes that each of the pitch events in the 13er group appears to have one main frequency region that is represented by orange-red and at least two overtones. The first overtone frequency region appears parallel over the main frequency in a lighter yellow. The second overtone frequency region parallel above the first and second frequency regions appears in green. The 13er group pitch events sustain, with the help of reverberation, through the attack of new notes, and add frequencies to the spectrum of consecutive pitch events, thereby blurring their timbral identities and creating a timbre that is inharmonic in character. In most cases, the spectrum of a previous pitch event is sustained throughout the attack and sustain of subsequent pitches.

The 14th pitch event consists of a main frequency region and at least 7 accompanying overtones of differing intensity levels. The lowest note consists of

[81] In order to account for all the pitches, this excerpt has been mixed down from 4 tracks to one mono track.

the darkest red, and hovers over an entire breadth of frequencies. The following overtones diminish in intensity levels, changing from red to orange to yellow to green, as the frequency increases and the breadth of frequencies diminishes.

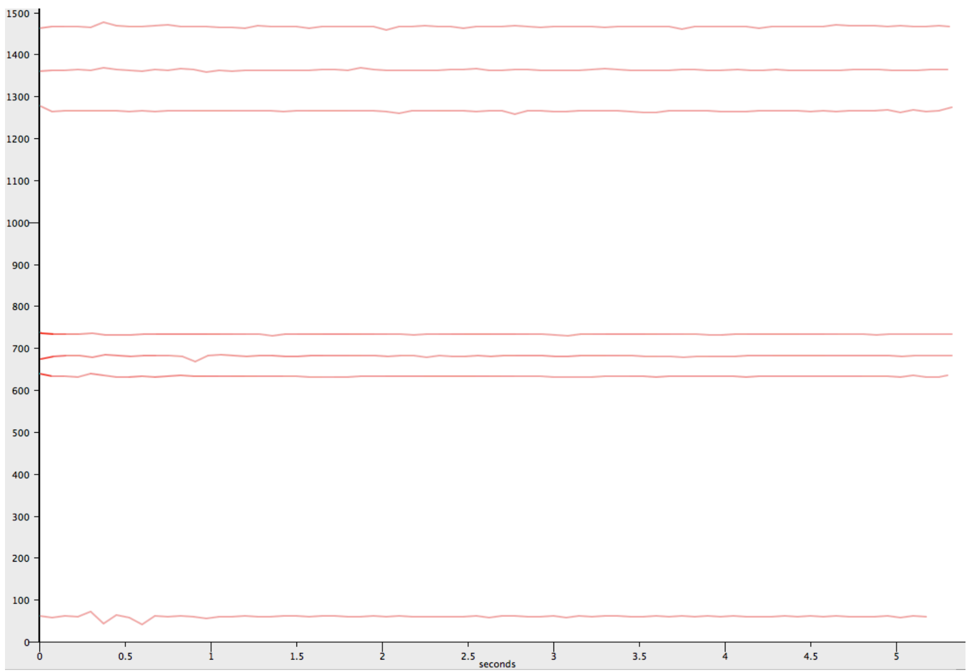

Figure 6: Spectrogram of peak frequencies.

Figure 6 shows a spectrogram that emphasizes frequency peaks. These frequency peaks are heard as a chord.[82] They are indicated in Hertz on the y-axis. The x-axis is phenomenal time in seconds. The period of the x-axis is about 5 seconds in duration. Lines of various shades of red represent the pitches. The shades of red indicate the intensity levels of the corresponding pitches. The darker the red is, the higher the intensity level of the pitch is.

Event	Range (Hz)	Pitch	PC	Cent Min	Cent Max
1	392	G3	7	1	1

[82] Figure 6 was generated with *SPEAR*.

Event	Range (Hz)	Pitch	PC	Cent Min	Cent Max
2	732-737	F#4	6	-33	-19
3	678-683	F4	5	-39	-24
4	484-489	B3	11	-17	4
5	705-710	F4	5	16	19
6	759-764	F#4/G4	6 to 7	44	-45
7	398-403	G3	7	26	48
8	608-613	D#4	3	-40	-26
9	845-850	G#4	8	30	40
10	545-549	C#4	1	-36	-17
11	818-823	G#4	8	-26	-16
12	915-920	A#4	10	-32	-23
13	586-592	D4	2	-4	14
14a	629-635	D#4	3	19	35
14b	678-683	E4/F4	4 to 5	49	-39
14c	732-737	F#4	6	-19	-7

Table 2: Opening pitch events (13er group and 14th pitch event).

Table 2 shows the pitch events of the 13er series and the compound pitch group of the 14th pitch event. The second column shows what the range of frequencies is in Hz. A range is indicated because Chowning builds little irregularities or fluctuations into each pitch.[83] The third column indicates the "traditional" pitch names. The 4th and 5th column indicates how these pitch events deviate in cents from the actual named pitches. If an entry shows two pitches then the 4th column will show the deviation of the first named pitch and the 5th column shows the deviation of the second named pitch. This occurs if the

[83] Chowning explains the use of little irregularities and fluctuations, or random vibrato, at the attack and sustain levels of pitches in his article: Chowning, "Frequency Modulation Synthesis of the Singing Voice." Here he applies the same principle to a timbre that is metallic in nature.

frequency range falls in between two pitches or quartertone as is the case with pitch events 6 and 14b.

Observing the pitch content, as illustrated in Table 2, it becomes immediately apparent that the composition utilizes a tuning system that is based on the harmonic series since all pitches, with the exception of the first one is detuned by varying degrees of cents.[84] For example, pitch 14b could either be an E or an F, but considering that the 11th partial of the harmonic series is about 49 cents flat, it is probably an F. Dividing the value of 680Hz by 11 results in Bb or 60Hz shown Table 3 in partial position 1. Table 3 shows a tetrachord with the pitch class set {3,5,6,T} that belongs to set class (0237):

Partial position	**Pitches in Hz**	**Pitch Class**	**Pitch Name**	**Cents**
1	59	10	Bb-1	20
2	632	3	Eb-5	25
3	681	5	F-5	-46
4	733	6	F#-5	-17
5	1265	3	Eb-6	28
6	1363	5	F-6	-43
7	1467	6	F#-6	-16

Table 3: Pitch event 14 displaying set class 4-14 (0237).

This tetrachord is produced by pitch event 14. Thereby, the composition's opening gestures sustain Chowning's aesthetic of creating music with the computer "that could not possibly exist in the 'real' world."[85]

[84] Chowning's previous composition *Stria* did utilize an alternate tuning system based upon an octave division through the Golden Mean (instead of 2:1 he used 1.618:1).
[85] Chowning, "DVD Program Notes," 109.

The first 7 pitch events present a slightly detuned offset palindrome with pitch class 11 functioning as a mirror point. The palindrome is obfuscated by the accelerating rhythmic nature of the 13er group. The pitch classes in the palindrome read 7, 6, 5, 11, 5, 6 and 7, where pitch classes 5, 6 and 7 are on average tuned about a quartertone sharp. The next set of pitch classes reads 3, 8, 1, 8 and one would expect the next pitch class to be 3. However, Chowning inserts pitch class 10 and 2. Pitch class 3 does appear as the bottom note of the following 14th pitch event.

Because of the distinct nature of each individual pitch in the 13er group that ends with pitch class 4-14 on pitch event 14, it is unmistakable that Chowning serialized 16 distinctively different pitches within the range of an octave plus a slightly flat tritone. Table 4 shows the same pitches sorted from the lowest to highest pitches. Sometimes the frequency ranges may intersect, but the rates at which the pitch events are sharp or flat in cents shows that the pitches are intentionally different, even if they seem identical; in the same way a singer's voice might sing the same pitch twice, but each time it may be slightly different, when analyzed through spectral analysis.[86]

Event	Range (Hz)	Pitch	PC	Cent Min	Cent Max
1	392	G3	7	1	1
7	398-403	G3	7	26	48
4	484-489	B3	11	-17	4
10	545-549	C#4	1	-36	-17

[86] Chowning, "Frequency Modulation Synthesis of the Singing Voice."

Event	Range (Hz)	Pitch	PC	Cent Min	Cent Max
13	586-592	D4	2	-4	14
8	608-613	D#4	3	-40	-26
14a	629-635	D#4	3	19	35
3	678-683	F4	5	-39	-24
14b	678-683	E4/F4	4 to 5	49	-39
5	705-710	F4	5	16	19
2	732-737	F#4	6	-33	-19
14c	732-737	F#4	6	-19	-7
6	759-764	F#4/G4	6 to 7	44	-45
11	818-823	G#4	8	-26	-16
9	845-850	G#4	8	30	40
12	915-920	A#4	10	-32	-23

Table 4: As Table 2, but now ordered from low to high.

The 14th pitch event contains pitch class 6 as the top part of the 3-note cluster. However, the 3-note cluster that consists of pitch class set {3,5,6} is in fact a subset of pitch class set {3,5,6,T} (Table 3). The top 3 note's central pitch dissects the distance of pitch class 3 and pitch class 6 in half and sits right in between pitch class 4 and 5. However, as previously mentioned, it is really pitch class 5 because it is the 11th partial of 60Hz and the 11th partial tends to be a quarter tone sharp. From the configuration of set class (0237), a beating effect is built into pitch event 14 that simulates the acoustic properties of inharmonic frequencies fading in and out.[87] The mind perceptually fuses tetrachord 4-14 into a unique metallic timbre, a phenomenon that Chowning refers to as perceptual

[87] Timbres composed of harmonic frequencies are multiples of integers, similar to string or wind instruments, whereas timbres constructed out of inharmonic frequencies are multiples of floating point numbers, similar to percussion instruments.

fusion.[88] This tetrachord has now lost its function as a chord and has become a ringing bell timbre. Chowning uses a technique here that Risset employs in the composition Mutations from 1969, because the concept is “uniquely possible with computers.”[89] The technique links “timbre to pitch space.”[90] Chowning reports:

> He [Risset] composed a short pitch sequence that is heard first sequentially in time (melody), then simultaneous in time (harmony), and then again simultaneously with exactly the same pitches but now as partials associated with a single sound source…. Because all of the partials die away in a similar manner, they fuse and are heard as timbre rather than harmony.[91]

Table 2 also functions as a roadmap with which to further analyze the timbre with *SPEAR* and how the spectral envelope was synthesized below the surface level. *SPEAR* is a sinusoidal partial editing, analysis and resynthesis tool that draws each partial of the spectral envelope of each sound that occurs. It has the advantage of revealing what type of synthesis method may have been used in the creation of the timbre of each of the pitch events. Tracking partials with *SPEAR* further underlines the necessity of Table 2. All partials are highlighted in red in Figure 7.

---

[88] Jean-Claude Risset, "The Perception of Musical Sound." http://www.utexas.edu/cola/insts/france-ut/_files/pdf/resources/risset.pdf.
[89] Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future," 5.
[90] Ibid.
[91] Ibid.

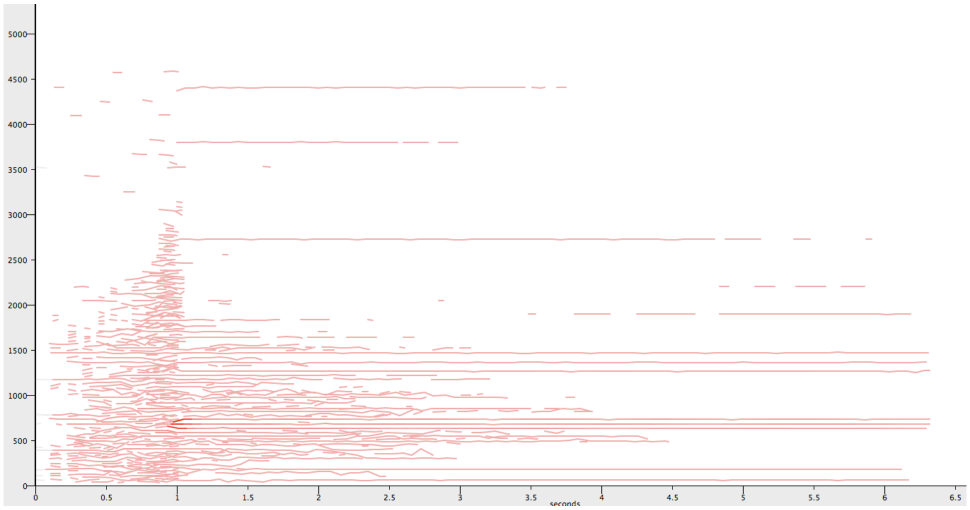

Figure 7: Analysis of the same 6.5 seconds.

Zooming in on the first pitch event of the 13er group (Figure 8) other peak frequencies of the spectral envelope that were perceptually fused or were previously indistinguishable, become evident. This pitch event only lasts 0.087 seconds.

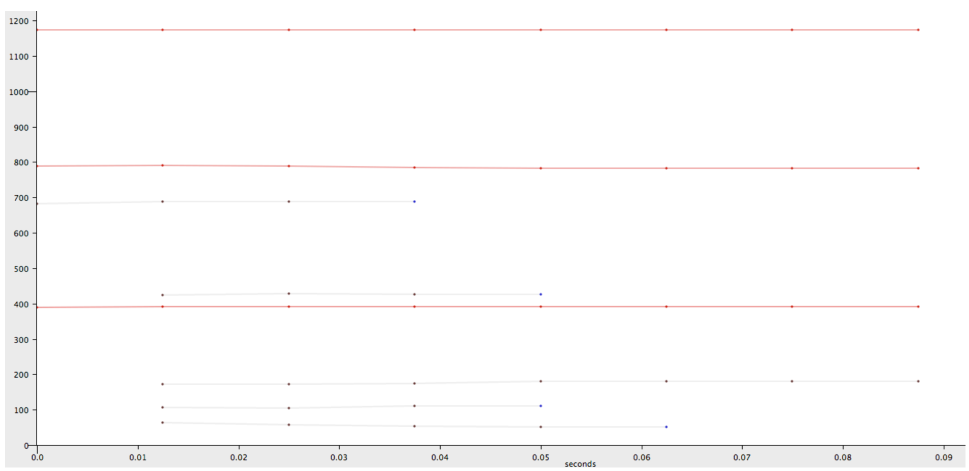

Figure 8: Analysis of the first pitch event of the 13er group.

The first partial is based on the 2:1 ratio, one perfect octave above the fundamental frequency. The second partial is based on the 3:2 ratio which is situated one perfect octave plus a perfect fifth above the fundamental frequency. The information of the precise values of the fundamental plus 2 partials (red lines) extracted with *SPEAR* is shown in Table 5.

Partial	**Start Time (seconds)**	**Frequency (Hz)**	**Intensity (out of 1)**
Fundamental	0.0000	389.87	0.00208
(1:1)	0.0125	390.41	0.00273
	0.0250	390.72	0.00287
Average Hz:	0.0375	390.90	0.00274
390.75	0.0500	390.97	0.00261
	0.0625	391.00	0.00253
	0.0750	391.05	0.00243
	0.0875	391.08	0.00233
1st Partial	0.0000	789.34	0.00007
(2:1)	0.0125	789.59	0.00007
	0.0250	787.65	0.00007
Average Hz:	0.0375	783.78	0.00008
784.82	0.0500	781.90	0.00011
	0.0625	782.03	0.00013
	0.0750	782.07	0.00015
	0.0875	782.17	0.00018
2nd Partial	0.0000	1173.79	0.00053
(3:2)	0.0125	1173.41	0.00069
	0.0250	1173.16	0.00072
Average Hz:	0.0375	1172.98	0.00069
1173.19	0.0500	1172.89	0.00066
	0.0625	1172.97	0.00064
	0.0750	1173.10	0.00062
	0.0875	1173.17	0.00060

Table 5: Partials of the first note of the 13er group.

This compound pitch can be easily recreated via additive synthesis in *Pd* by utilizing Table 5 and converting it to a `qlist` (Table 6).[92] The header (e.g.: "1.294 step 1;") directs how long the duration of a specific pitch level is to last (usually 1.2494 in 1/1000th of a second), how the partials in Hz are changing, and how much of an intensity level can be expected for the individual partial. Because the intensity levels were measured, an ADSR envelope is not needed.[93]

```
step       0;
partial01  0;        Intensity01  0;
partial02  0;        Intensity02  0;
partial03  0;        Intensity03  0;
   1.2494  step      1;
partial01  389.87;   Intensity01  0.00208;
partial02  789.34;   Intensity02  0.00007;
partial03  1173.79;  Intensity03  0.00053;
   1.2495  step      2;
partial01  390.41;   Intensity01  0.00273;
partial02  789.59;   Intensity02  0.00007;
partial03  1173.41;  Intensity03  0.00069;
   1.2494  step      3;
```

[92] A qlist in Pd is an external text file that can read timed events into a given Pd-Patch. This text file acts as a two-dimensional array or database-like table. Each step contains how long the note lasts (array1[step[1…8]] in **bold**), and then maps to an array that contains the partials and their corresponding intensity levels (array1[array2[partial[1…3], intensity[1…3]]). Pd itself cannot map the duration values of each individual step (0.012494 seconds), but the `qlist` can be read one step at a time, which allows for close examination of the individual 8 changes of timbre that occur within the first note of the 13er group. However, if all the different voices were mapped to a `poly` object the pitches would play successively. How polyphony can be used in Pd is illustrated in Figure 42 in APPENDIX B – PD PATCHES.

Dodge and Jerse define additive synthesis as a "production of sound by direct summation of component frequencies…each component is produced by a separate sinusoidal oscillator." Dodge, 429.

[93] "An envelope generator…makes an audio signal that smoothly rises and falls as if to control the loudness of a musical note.… 'ADSR' is an acronym for 'Attack, Decay, Sustain, Release', the four segments of the ADSR generator's output." Miller Puckette, *The Theory and Technique of Electronic Music*, 89-90.

```
partial01  390.72;   Intensity01  0.00287;
partial02  787.65;   Intensity02  0.00007;
partial03  1173.16;  Intensity03  0.00072;
   1.2494  step      4;
partial01  390.90;   Intensity01  0.00274;
partial02  783.78;   Intensity02  0.00008;
partial03  1172.98;  Intensity03  0.00069;
   1.2495  step      5;
partial01  390.97;   Intensity01  0.00261;
partial02  781.90;   Intensity02  0.00011;
partial03  1172.89;  Intensity03  0.00066;
   1.2494  step      6;
partial01  391.00;   Intensity01  0.00253;
partial02  782.03;   Intensity02  0.00013;
partial03  1172.97;  Intensity03  0.00064;
   1.2494  step      7;
partial01  391.05;   Intensity01  0.00243;
partial02  782.07;   Intensity02  0.00015;
partial03  1173.10;  Intensity03  0.00062;
   1.2494  step      8;
partial01  391.08;   Intensity01  0.00233;
partial02  782.17;   Intensity02  0.00018;
partial03  1173.17;  Intensity03  0.00060;
```

Table 6: A *Pd* style cue list (`qlist`).

The following figures show how additive synthesis can be implemented with the preceding `qlist` in *Pd* and how to step through the individual 8 time slices in order to hear the minimal transformation of timbre of the first pitch event of the 13er group. Building the patch gives a clue to how the actual computer music technique is viewed and implemented as a music theory entity and contributes to the overall reconstruction of X at the end of this study.

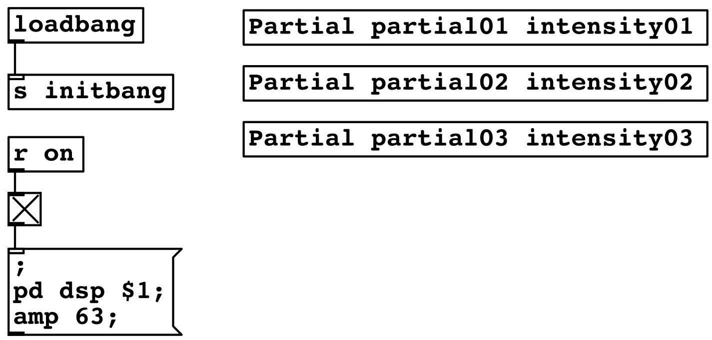

Figure 9: Initialization Function.

Figure 9 shows an object containing a function that runs at load time of the patch.[94] This function prepares 3 objects (Figure 10 shows the object) consisting of 3 oscillators each. Further the 3 oscillators contain the 3 partials and their corresponding intensity levels (e.g.: `Partial partial01 intensity01`). The initialization function also makes sure that audio is computed in Pd (`pd dsp $1;`) and specifies a volume level (`amp 63;`). This function also has a built in indicator that can be used for other functions signifying that the patch is loading

[94] Initialization functions are used in all types of programming languages, e.g.: C, C++, Java, and even scripting languages like JavaScript and its relative Actionscript. Its use is considered good programming practice. For example, an initialization script in Objective-C could look like the following small program:

```
#import<Foundation/Foundation.h>

int main (int argc, const char * argv[])
{
      NSAutoreleasePool * pool = [[NSAutoreleasePool alloc] init];
      NSLog (@"Foo!");

      [pool drain];
      return 0;
}
```

(Stephen G. Kochan, *Programming in Objective-C 2.0* (Upper Saddle River, NJ: Addison-Wesley, 2009), 9.)

via the loadbang object, in other words it sends a bang to wherever specified (s initbang).

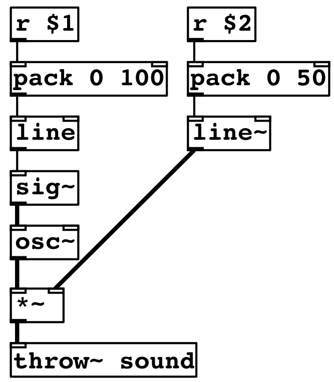

Figure 10: Oscillator for Additive Synthesis.

Figure 10 shows how an assigned Hz value (r $1) and an intensity value (r $2) are used to create an oscillator that is used for the additive synthesis process.[95] $1 value sends its parameter to the left outlet of the osc~ object and is then multiplied via the *~ object with the $2 value. The newly created parameter then is added to a sum bus via the throw~ object.[96] The line and line~ objects are used to ensure a smooth timed transition in order to avoid clicks that result from rapidly changing values.

[95] A $ symbol is used to access an assigned variable, similar to the $ token in the PERL or PHP programming languages. $1 (e.g.: partial01) and $2 (e.g.: intensity01) were set in the creation argument of the partial object in the initialization function.
[96] The sum bus is the product of the added partials and contains the combined additive synthesis.

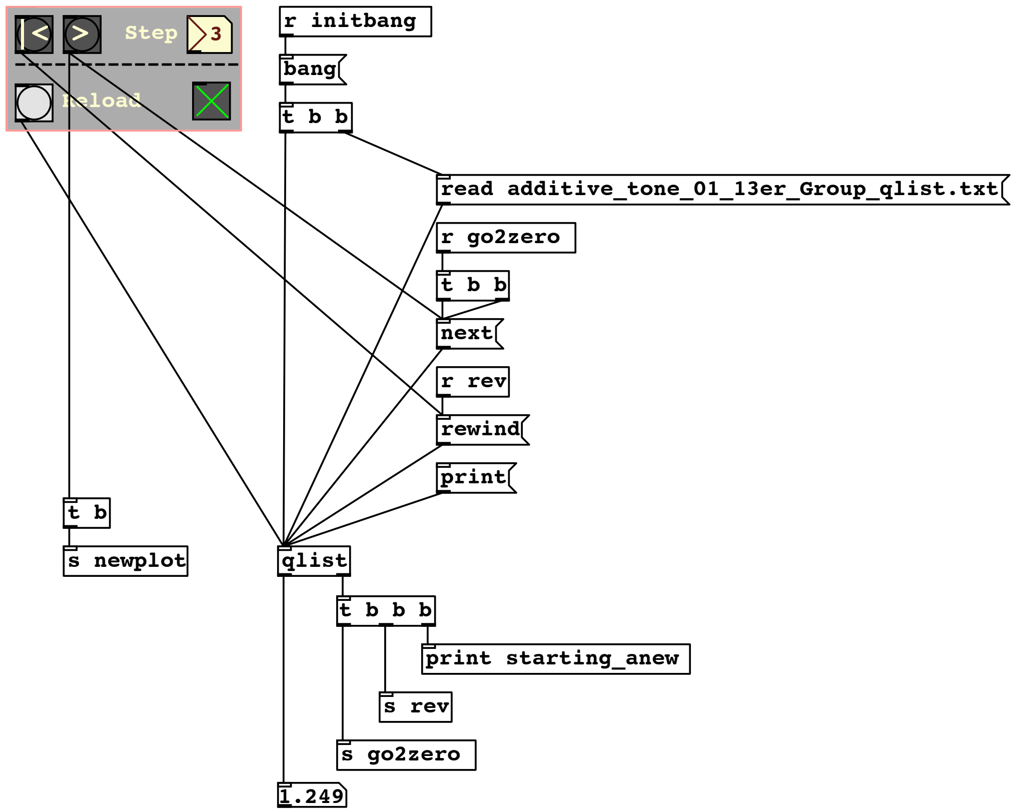

Figure 11: Loading, routing and navigating the `qlist`.

Figure 11 shows the object (`Additive_Tone_Stepper.pd`) that deals with loading the data arrays from the `qlist` and routes them to the previously specified creation arguments in the `partial` objects. The object from Figure 11 gets an `initbang` from the `init` sub patch, which reads the `qlist` from Table 6 and automatically runs this list one time through the specified timed values. Once the list of events has completed, a series of `bang`s are sent in sequential order via the `trigger` (`t`) object. The first `bang` sends a message to the Pd window to signal that the `qlist` will begin anew (`print starting_anew`). The next bang is sent to the rewind message and rewinds the `qlist`. To reset the qlist to a standby step before the 1st step the final bang is send to the next message (go2zero).

The qlist can be reloaded (bang) and rewound (rewind) to the standby step any time via two labeled bangs in the user interface. The individual timed events can also be stepped through one at a time with the next message that is activated with a bang in the user interface. The bang that steps through the timed events, is also used to send a bang to the upcoming fft object in order to re-plot the partials in the spectrum chart for each individual event (s newplot).[97] A print message is supplied to the qlist object for debugging purposes in the Pd window, i.e. did the qlist load and what parameters are contained within the qlist. The number2 (e.g.: Step 3) object gets the parameter specified in the qlist that shows what sequential step is currently being synthesized. The user interface canvas also controls a mechanism via a toggle button (a squared gray object next to the Reload bang) that can turn audio computing on or off at any given time.

---

[97] FFT stands for Fast Fourier Transfer analysis.

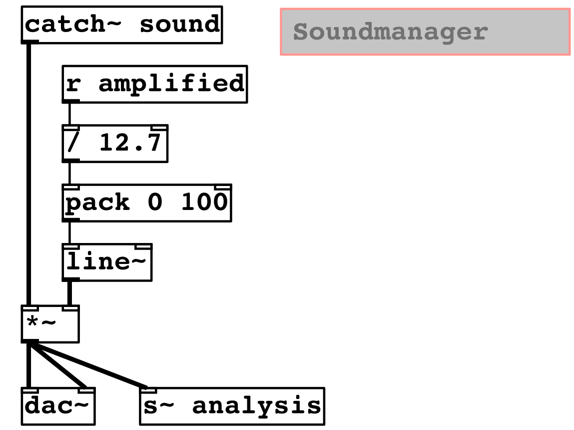

Figure 12: The sound manager object.

Figure 12 shows the `soundmanager` object. Its purpose is to `catch~` the sounds that the `throw~` object from the `partial` object added to the sum bus and sends the combined sound to the digital audio converter object (`dac~`), which produces sound on a speaker system attached to the computer. Since the first sound event of the 13er group contains a faint signal, an amplifier was provided that multiplies the caught sum bus, in order to be able to hear the event clearly. The multiplied signal is also sent to the `Plot_Spectrum.pd` object (`s~ analysis`).

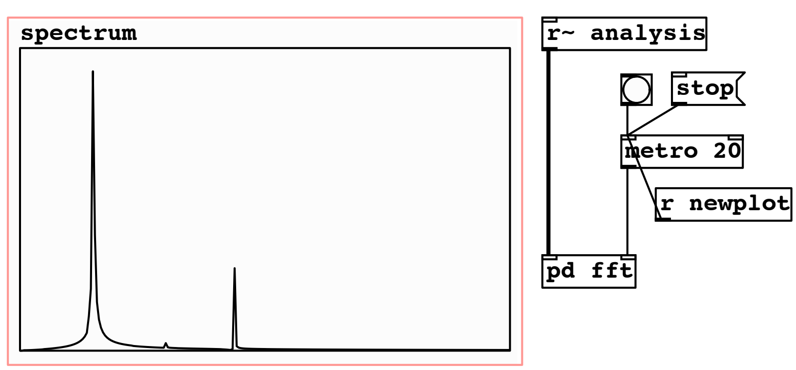

Figure 13: Object for spectrum calculation.

Figure 13 shows the spectrum object. This object receives a readied audio signal and sends it to the `fft` analysis object (Figure 14). Once the `fft` object has evaluated the audio signal, it is sent to an array that plots the `spectrum` of the audio signal in graph format. The y-axis indicates the intensity level and the x-axis shows the harmonic partials within the spectral envelope. There is also a function that receives a command when to plot a new graph (`r newplot`). Manual options for the metro object are also provided. The `bang`, `stop` message and the `metro` object with a creation argument (`20` milliseconds) that specifies how fast the `metro` object is to refresh and send another `bang` to the `fft` object that is located in the `pd fft` subpatch.

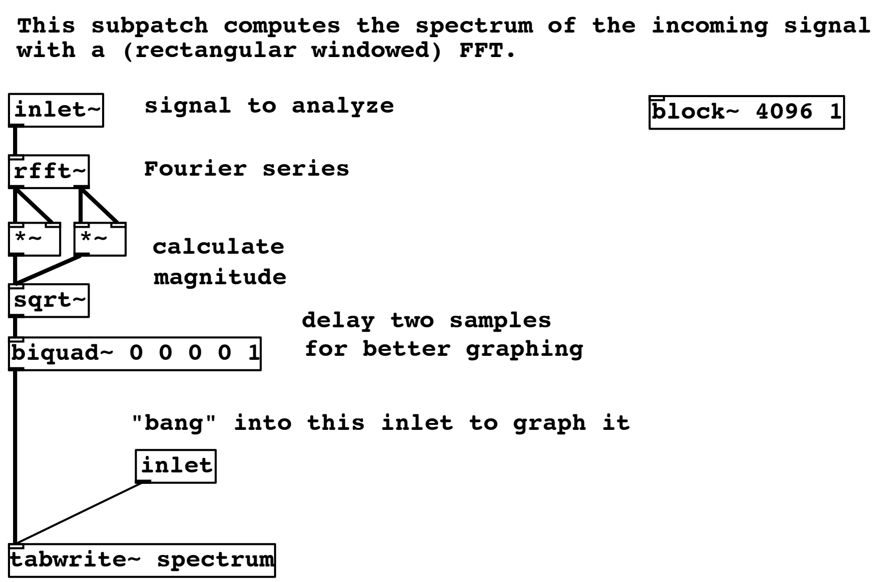

Figure 14: FFT subpatch, directly taken from the Pd documentation.[98]

[98] For further information on why and how this patch works consult Chapter 9 in Puckette, *The Theory and Technique of Electronic Music*, 267-299.

Figure 14 shows how the `fft` subpatch works. Its `inlet~` receives the audio signal from `r~ analysis` and performs the FFT analysis via the `rfft~` object.[99] The Fourier series is then squared and square-rooted in order to calculate its magnitude and delayed by two samples for better graphing. The `block~` size indicates the local sampling rate and specifies the overlapping computations.[100] The computed FFT parameters are then fed into the spectrum array (`tabwrite~ spectrum`) which requires a `bang` to be plotted.

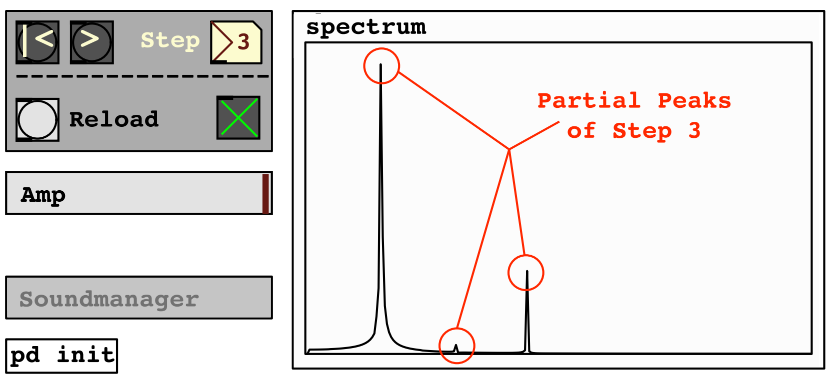

Figure 15: The main patch.

Figure 15 demonstrates how all the previous patches and subpatches (or objects) are integrated into a GUI (Graphic User Interface) which in itself is an object. The philosophy behind this type of programming is to be able to reuse previous code as often as possible, rather than re-inventing the wheel every time. This practice is also known as DRY (Don't Repeat Yourself). Every *Pd*

---

[99] The `rfft~` object is used for real fast Fourier transform analysis. More on this object and why it is more efficient than the `fft~` object can be found at Ibid., 291.
[100] Ibid., 214.

programmer/composer will have a library (or a collection of objects) at his or her disposal. Miller Puckette's book *The Theory and Technique of Electronic Music* is in itself his collection of objects to solve common computer music programming problems.[101]

The purpose of this patch is to be able to step through the 8 events within the first note of the 13er group. When using this patch the changes in intensity levels of the partials are reflected in the graph. These alterations mostly encompass the changes that happen within an ADSR envelope. Surprisingly this patch also shows how the timbre already changes over time within the 0.09 seconds of the first pitch event. The pitch seems to be rising in its decay phase and thereby contributes to a very slight timbral transformation. Timbral transformations are at the heart of the X-object. However, it also shows that additive synthesis is cumbersome and not very efficient from a computing standpoint, because too many parameters are required to re-synthesize a sound. A more effective solution is FM synthesis.[102]

For the discussed 13er group Chowning probably used a relatively simple FM synthesis configuration. The proposal here would be at least a double carrier configuration similar to algorithm 3 of the TX81Z as mentioned in the STK toolkit

---

[101] Objects also exist in music theory, and are combinations of music theory entities as described in footnote 62.
[102] However, FM synthesis is not as stable and only after much experimentation can the same goal of reproducing a sound be achieved.

and the ChucK manual, which Cook and Wang lovingly call “heavy metal.”[103] The C-Sound manual also contains this algorithm, calls it *fmmetal*, and makes the distinction that it uses 4 oscillators in the TX81Z configuration.[104] Algorithm 3 is just an extension of the FM synthesis base class i.e. it adds functionality on top of the regular FM algorithm in the STK toolkit.[105] Chowning describes how to assemble an FM synthesis algorithm for a bell-like sound.[106] He sets forth 2 premises, “1) the spectral components are not usually in the harmonic series [inharmonic], 2) the evolution of the spectrum is from the complex to the simple.”[107] Chowning further elucidates, “Bell-like sounds can be produced by making the change of the [modulation] index directly proportional to the amplitude envelope.”[108] He describes an exponentially decaying envelope that can be built like Figure 16.

---

[103] Perry R. Cook and Ge Wang, "The Chuck Manual," (Princeton: Princeton University, 2007). http://chuck.cs.princeton.edu/release/files/chuck_manual.pdf.
[104] Barry Vercoe, *The Canonical Csound Reference Manual: Version 5.10* (MIT Media Lab), 776.
[105] Cook and Wang, "The Chuck Manual."
[106] John M. Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," *Journal of the Audio Engineering Society* 21, no. 7 (1973): 533.
[107] Ibid.
[108] Ibid.

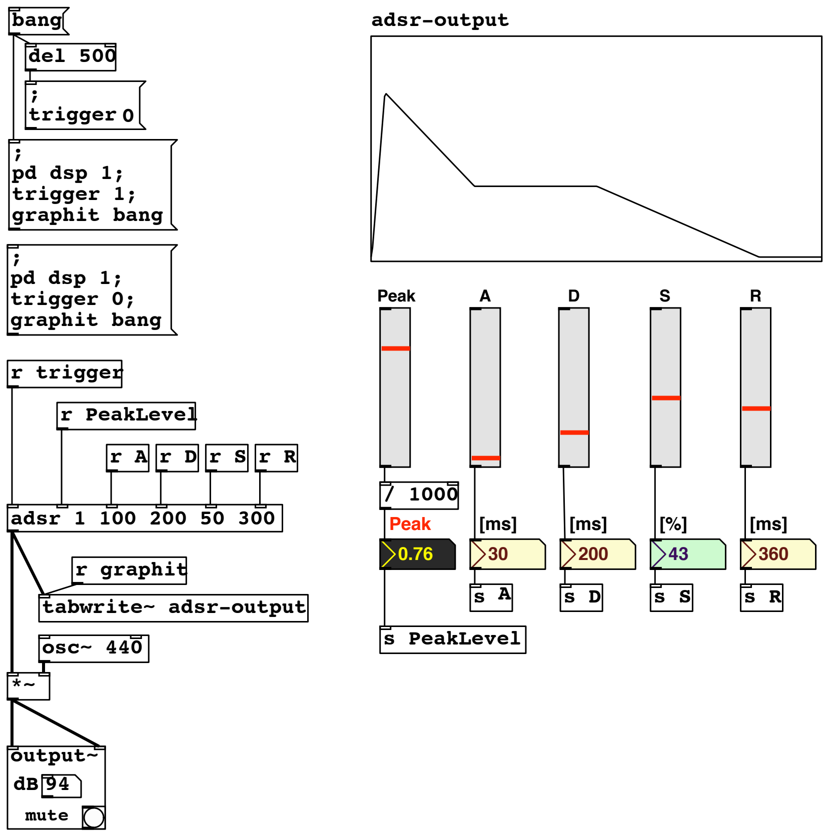

Figure 16: Envelope Generator, displaying an Intensity envelope of a bell.[109]

SPEAR provides more, accurate information about the pitch material than *Sonic Visualizer*. Because *SPEAR* provides more information, it becomes more difficult to isolate individual pitches. However, if one isolates the highest partials of the 13er group a clearer picture emerges, since after evaluating the first pitch event of the 13er group revealed that partials are harmonic. An average of these upper partials of the pitch events is shown in Table 7.

[109] Based upon an example in Miller Puckette, *The Theory and Technique of Electronic Music*, 104-107.

With this information and the approximation of the pitches of the 13er group from Table 2 it is possible to calculate what partial number the upper partial is for each individual pitch event, by dividing the Hz number of the pitch event from Table 2 with the Hz number from the upmost partials of Table 7. Knowing what the partial number of the upmost partial number is, one can divide the upmost partial in Hz by the partial number. This operation yields what the carrier frequencies are.[110] Table 7 shows the result of this operation.

Event	Average	Partial	Frequency	PC	Tuning (cents)
1	3519	9	391	7	-5
2	4405	6	734	6	-14
3	4093	6	682	5	-41
4	3421	7	489	11	-19
5	4244	6	707	5	21
6	4569	6	762	6	49
7	3248	8	406	8	-40
8	3662	6	610	3	-34
9	4257	5	851	8	42
10	3821	7	546	1	-27
11	4104	5	821	8	-21
12	4581	5	916	10	-31
13	3525	6	587	2	0

Table 7: Recalculated and average values of Table 2.

As previously mentioned, the cents tuning deviation indicates that these pitches are harmonics. The composer isolated harmonics, and then reassembled all pitches as a chord. The isolated harmonics are spaced as multiples of a

[110] The carrier frequencies are the frequencies that are distinguished and perceived by the mind and later can be used as the carrier frequencies for an approximate FM re-synthesis.

positive integer. In effect, they are not equal tempered, or Pythagorean, or Mean-Tone tempered, or tuned by ratios of the Golden Mean (as Chowning did in *Stria*), but are the partials of the harmonic series. Upon reassembly of the partials as simultaneities, a chord sounds. Chowning applies a metallic bell to the chord and the chord becomes a timbre, a metallic timbre.

The gesture of linking timbre to pitch space is used at the beginning of the composition. It is revisited several other times in the course of the composition. X-5 has a preliminary gesture that is 2 seconds long and begins at 2:10. It does not use metal sounds, but rather disassembles timbre based on notes from a harmonic series generated by a voice. Another occurs at the onset of the B section at 5:06 and lasts until 5:15. In B the gesture is developed in two subsections. The first subsection lasts until 7:03 and features metallic bell timbres juxtaposed with vocal timbre. The second subsection from 7:03 to 8:06, features a development where synthesized vocal sounds are highlighted in various polyphonic arrangements. In C a prominent use of timbre to pitch space linking is found from 9:03 to 9:06. Here it is applied as a preliminary gesture to X-19 that is located before another contrapuntal vocal treatment at the end of C (9:36 to 11:18). The gesture does not reemerge in D (coda). Since X appears as a result of this gesture, it is intrinsically related to X.

X-1

X-1 is about 30 seconds in length, starts at ca. 15 seconds and ends at ca. 45 seconds at the beginning of the composition.[111] It is the first fully realized occurrence of X. Figure 17 shows the waveform of X-1.

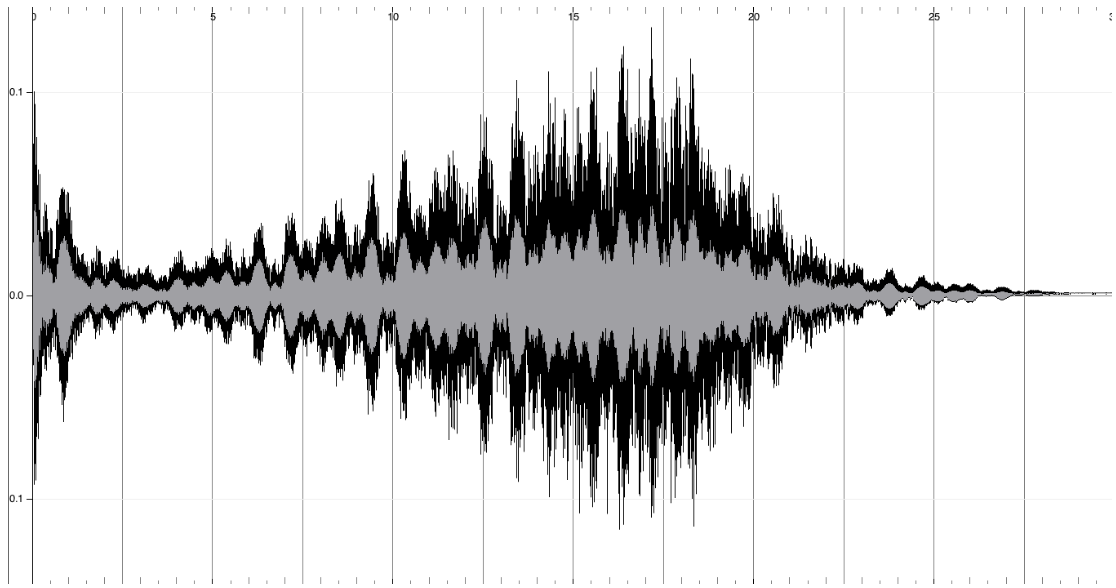

Figure 17: X-1 Waveform (mono).

The intensity envelope, measured from 0-1 in Figure 17, of X-1 maps onto the proposed X object. Object X-1 starts with a rapid attack and decays expediently, which is the head. The timbre is metallic, resembling that of a low bell. The timbral nature of object X-1's head was introduced in the previously discussed preliminary gesture, which consists of "heads" only.

The tail of object X-1 swells slowly and decays at about half the rate at which it gained intensity. The timbral content of the X-1 tail material is metallic in character. It is an expansion of the head material. The deep metallic bell sound

[111] The exact timeframe can be seen in the Events-Chart, Appendix A - Table 10.

from the head is expected to decay, but it is prolonged by the aforementioned parameters of an ADSR envelope. The procedure shows how Chowning is concerned with the manipulation of sounds that are only possible with the help of a computer. The tail does exhibit characteristics that are associated with low/deep vocal properties. Further analysis calls for the necessity of a spectrogram.

Figure 18 shows a spectrogram of X-1 as generated with *Sonic Visualizer*.[112]

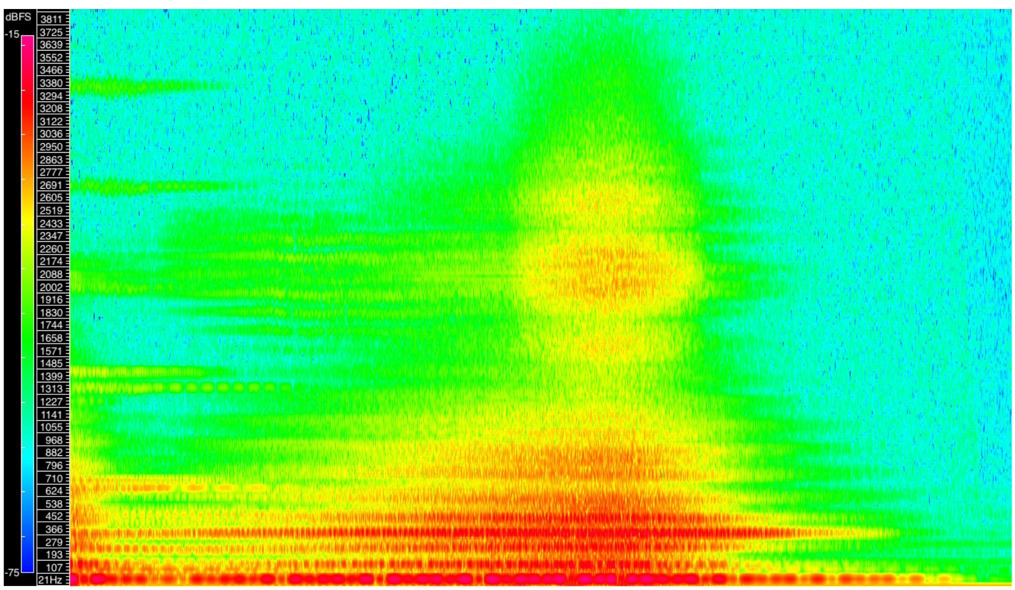

Figure 18: X-1 Spectrogram.

The red/purple color indicates the highest intensity levels of partials below 100 Hz. The x-axis indicates time in seconds and the y-axis shows increasing frequency levels in Hz or pitch range. The highest intensity levels are found with the lowest sounds. The head yields a few purple levels, ebbs off and continues

[112] Figure 17, Figure 18 and Figure 19 show the examples in mono output.

with increased intensity throughout its tail. Some mid range partials are detected around the 360 Hz range, with a few partials in between in the tail. Any partials above the 360 Hz range are decreasing in intensity levels and are displayed in an arch form. The head contains a range of partials reminiscent of the “heads” of the preliminary gesture and its attack ranges up to ca. 3380 Hz, indicative of the presence of formants.

The tail section of X-1 also shows partials of higher frequencies ranging from 1916 Hz to 2433 Hz. This elliptical shape mirrors one frequency range below from 1571 Hz to 1830 Hz (yellowish ellipse), and one frequency range above from 2519 Hz to 2777 Hz (another yellowish ellipse). Additionally, at the peak intensity level of the tail another formant region can be detected in the high register between 3380 Hz and 3811 Hz. The mirroring ellipses indicate the presence of simultaneously occurring notes that sound like a chord.

Most of the partials shown in this spectrogram (Figure 18) exhibit certain degrees of “waviness” indicative of an applied vibrato to each of the individual partials. These quasi-periodic vibrations are reminiscent of the vocal quality, but in the context of the metallic bell sound of the tail section. It is the random vibrato described in Chowning’s article *Frequency Modulation Synthesis of the Singing Voice*.[113]

---

[113] Chowning, "Frequency Modulation Synthesis of the Singing Voice," 62. Chowning provides the formula of vibrato percent deviation that equals to 0.2 x log(fundamental pitch frequency). The formula produces vibrato frequency ranges from 5 to 6.5 Hz.

Figure 19 shows a more detailed spectrogram of all partials contained in X-1 as generated with *SPEAR*. The previous observations are reflected within this analysis (now without the color-differentiated highlighting). The frequencies with the highest intensity levels are highlighted in a darker shade of red and reveal some insight into how the FM synthesis may have been configured to produce the sound mass of the tail of X-1.[114]

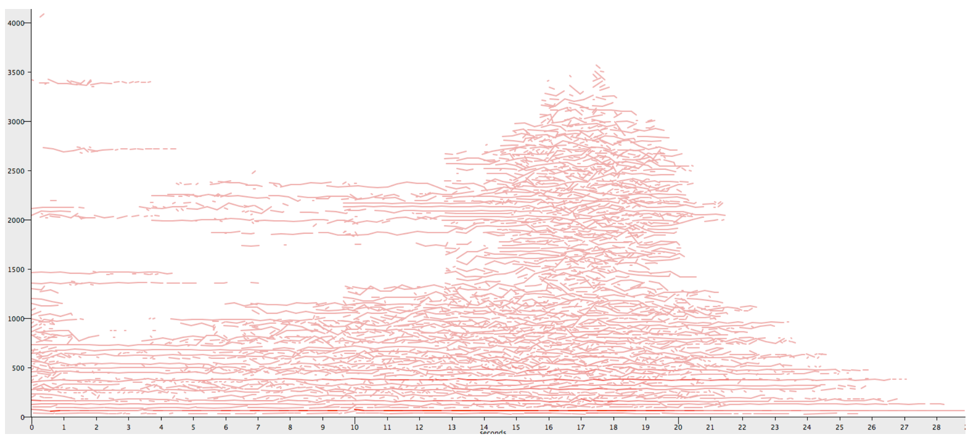

Figure 19: X-1 Partials.

Previously the observation was made that the low frequencies and their accompanying elliptical groups in the tail contain several simultaneous occurring sounds centered on several frequency groupings. The surrounding partials around these frequency groupings are inharmonic, as would be expected from a metallic sound. By reducing the threshold dB levels, it is possible to zoom in on some of the carrier frequencies that may be present, since they have the highest

[114] *Sound mass* describes simultaneously occurring pitch events, they can be chords just the spectral envelope of a compound timbre.

intensity levels (Figure 20). Table 8 shows an average of the estimated pitch values of the carrier frequencies.

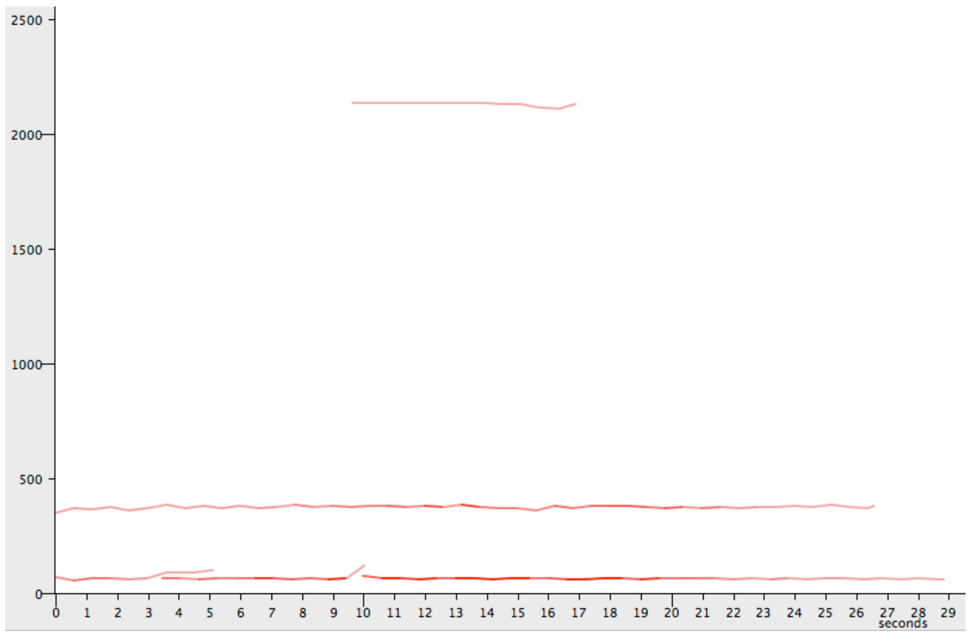

Figure 20: X-1 Tail Carrier Frequencies.

Carrier	Average Frequency (Hz)	Pitch Class	Deviation (Cents)
1	62.83	0	-43
2	373.73	6	17
3	2128.24	0	28

Table 8: Carrier Frequencies of X-1 tail.

A cross section at the peak level of all the partials in the tails section yields the following data (Figure 21):

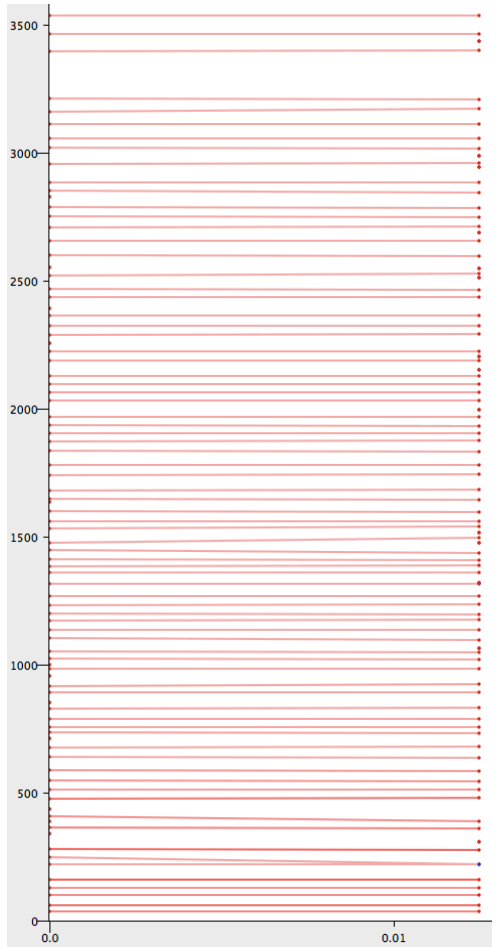

Figure 21: X-1 Tail Peak Partials.

From this data, it can be concluded that Chowning used several multi carrier FM generators. As proposed earlier the carrier frequencies indicate how many of the multi carrier FM generators were needed. The multi carrier FM generators were also responsible for inducing the varying degrees of random vibrato within the generated partials. Again, an FM algorithm similar to algorithm 3 (TX81Z) was used.[115]

---

[115] It should be mentioned that Chowning used an algorithm slightly more complex, i.e. he probably used more than just 4 oscillators to generate that sound.

X-3

Figure 22 shows the waveform of X-3 as generated with *Sonic Visualizer*. For simplicity, the sound file has been down-sampled to mono so that it can be viewed as one unit. What immediately jumps out at the observer is how X-3 maps onto the proposed X object. The X-3 object is about 5 seconds in length and is situated within A as shown in Table 10.

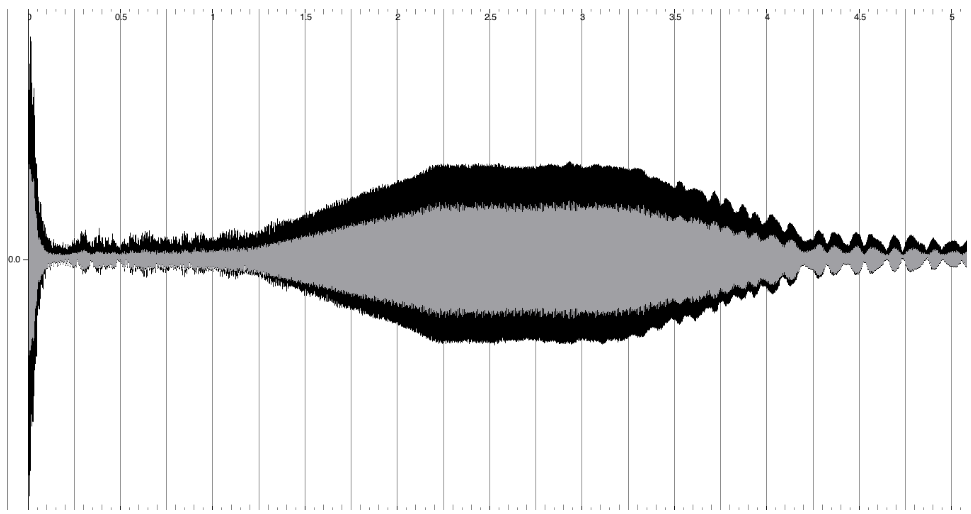

Figure 22: X-3 Waveform (mono).

The head of this object has a sharp attack and fast decay in its intensity envelope as has been previously observed with the "heads" of the preliminary gesture and the head of the X-1 object. The tail section swells and decays at equally proportioned levels in its intensity envelope.

However, the tail section now clearly creates the first truly audible vocal timbre. The vocal properties of the tail section are supported by the addition of random vibrato, which can be observed within the elliptical orb that outlines the wavy yellow partials of Figure 23 (see below). The spacing of partials is harmonic in nature and contains a strong fundamental (red) with supporting partials, which is supported by their equally spaced occurrences in the corresponding Hz ranges. The upper partials of the tail section, green in the ca. 4000-5000 Hz range, exhibit formant attributes that contribute to an audible vowel sound generated.

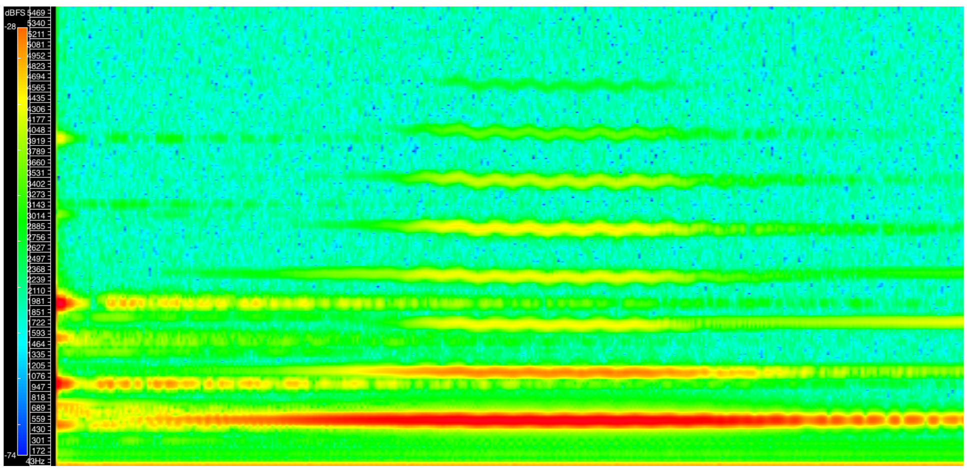

Figure 23: X-3 Spectrogram.

The head section shows a metallic sound with non-equidistantly spaced partials indicating the expected inharmonic character of the timbre. The reverberation applied to the head section of the decay phase of the intensity envelope traces to the tail section and the spectrogram. This is reflected by the X-3 spectrogram, because of the partials that fall in between the partials generated for the tail section. A more precise spectrogram of all the partials contained in X-3 as generated by *SPEAR* is shown in Figure 24.

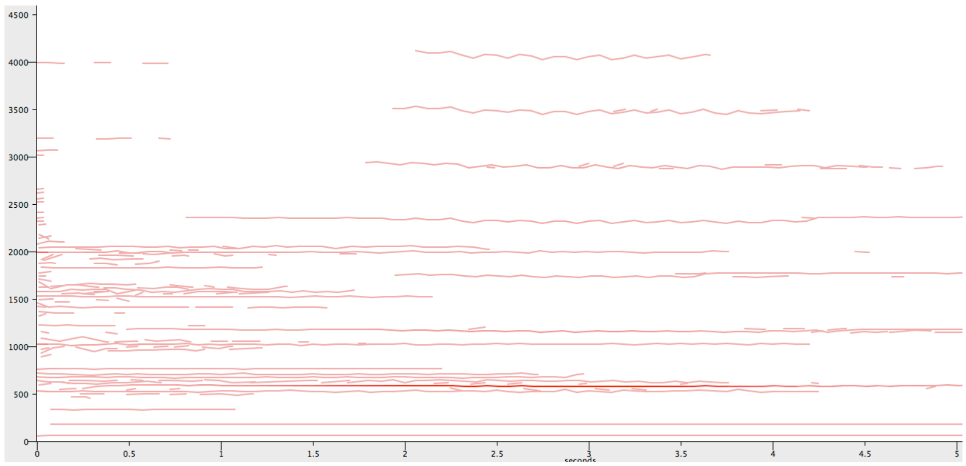

Figure 24: X-3 Partials.

However, the combined mono figures of X-3 are misleading. Examining this gesture one channel at a time yields a quite different result. Channel 4, for example, contains no data or only faint artifacts. Channel 1 and channel 2 (Figure 25) contain the metallic head of the object.

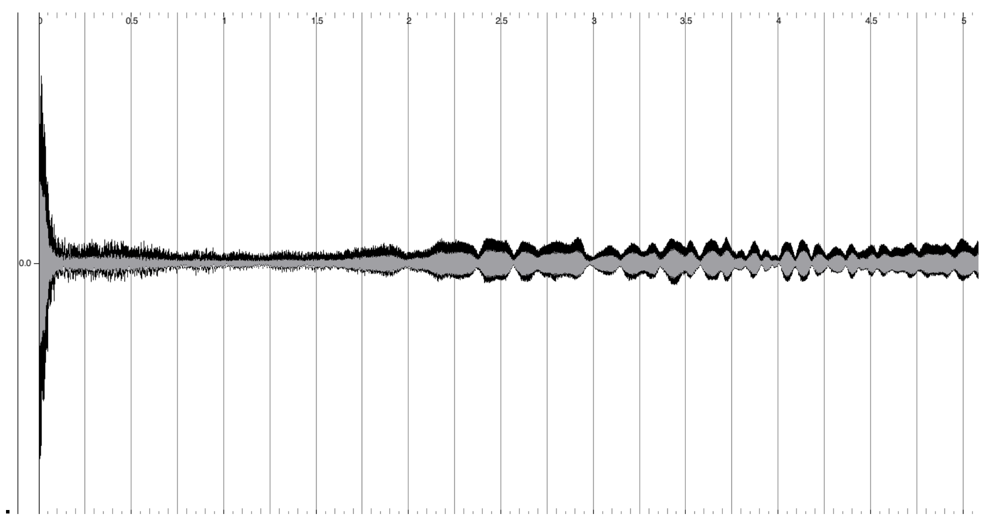

Figure 25: X-3 Channel 2 Waveform.

Channel 3 (Figure 26) contains the voice-like part of the synthesis, which slightly bleeds into channel 2.

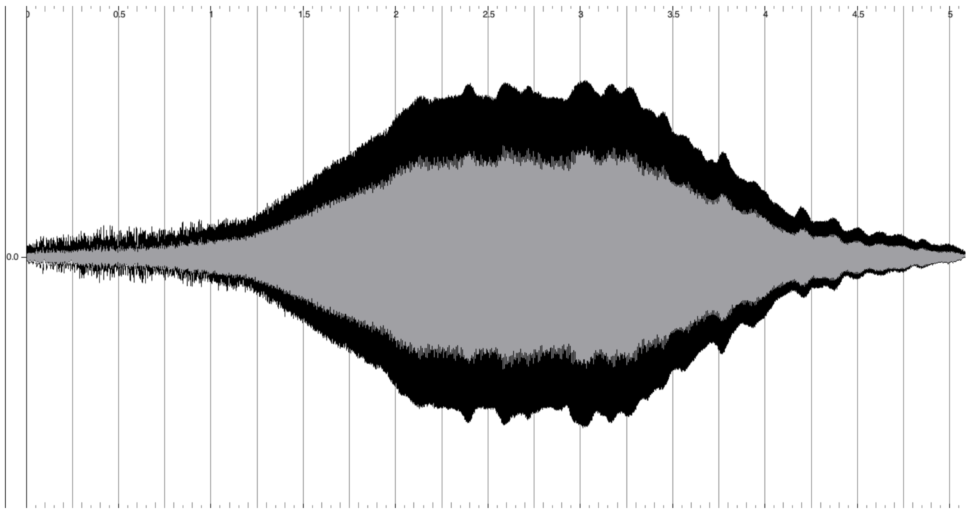

Figure 26: X-3 Channel 3 Waveform.

Separate partial analyses for the separate channels in this scenario will yield a clearer picture (Figure 27).

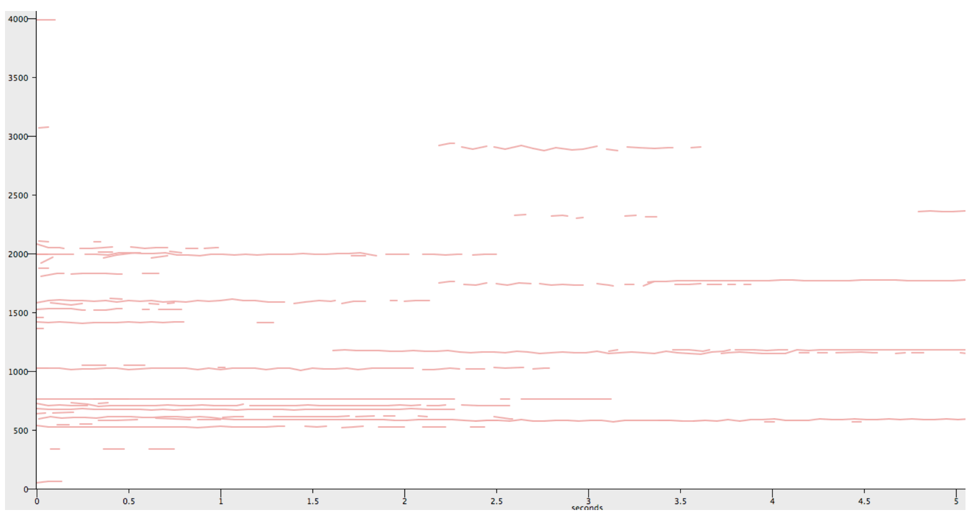

Figure 27: X-3 Channel 2 Head Partials.

A bisection of the head partials shows the initial inharmonic content of the attack phase of channels 1 and 2 and yields the following pitch content (Table 9).

An FM synthesis algorithm that is based on the bell-like tone described by Chowning in his article “The Synthesis of Complex Audio Spectra by Means of Frequency Modulation” generated the sonic material for both channels of the head section.[116] It is augmented by a variation of a multi carrier FM synthesis configuration.

Channel 1		Channel 2	
Intensity (0-1)	Pitch (Hz)	Intensity (0-1)	Pitch (Hz)
0	54.16	0	55.75
0.000134	531.11	0.000163	527.15
0.000323	594.21	0.002031	592.66
0.000122	640.86	0.00036	634.82
0.000124	675.71	0.000185	675.65
0.000232	722.93	0.000394	711.77
0.000223	761.75	0.000466	755.27
0.000431	1024.63	0.00062	1,024.51
0.005056	1365.79	0.006419	1,223.57
0.000131	1420.76	0.000239	1,348.67
0.000131	1459.92	0.000163	1,424.14
0.000114	1525.57	0.000158	1,460.07
0.000164	1581.25	0.000144	1,529.14
0.000499	1874.60	0.000288	1,579.61
0.000282	1994.76	0.001684	1,624.08
0.008528	2078.79	0.000096	1,675.94
0.000258	2106.11	0.000072	1,875.17
0.000251	3070.62	0.000344	1,994.80
0.000044	3989.32	0.011381	2,036.91
		0.001712	2,077.60
		0.000328	2,179.62
		0.000152	3,068.18
		0.000062	3,989.30

Table 9: X-3 Head Bisection Pitch Content.

Figure 28 shows the partials contained in the X-3 tail section. The waviness of the lines shows the vibrato.

[116] Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," 533.

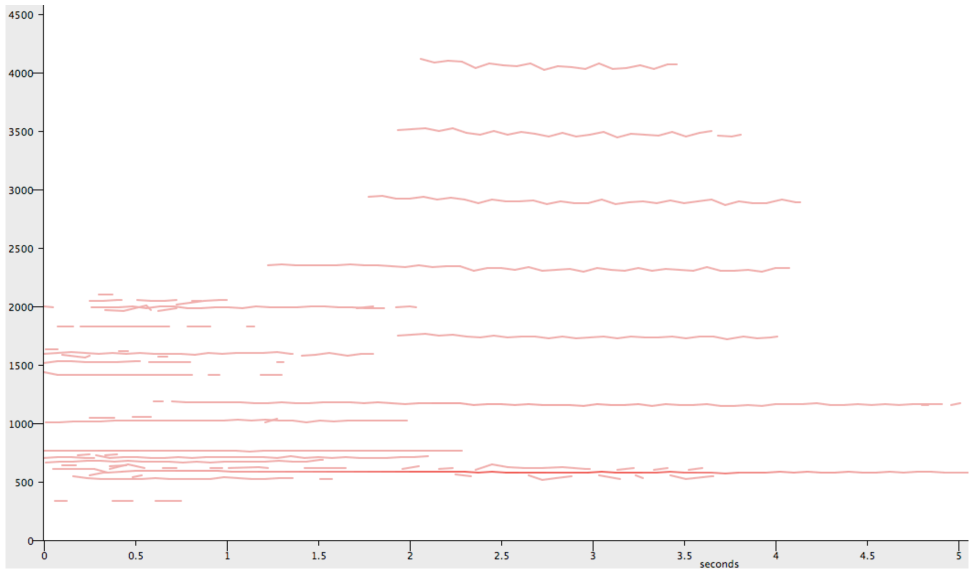

Figure 28: X-3 Channel 3 Tail Partials.

By removing the partials that bleed into channel 3 and only highlighting the carrier frequencies needed for the vocal synthesis yields a clearer picture of X-3's tail section (Figure 29).

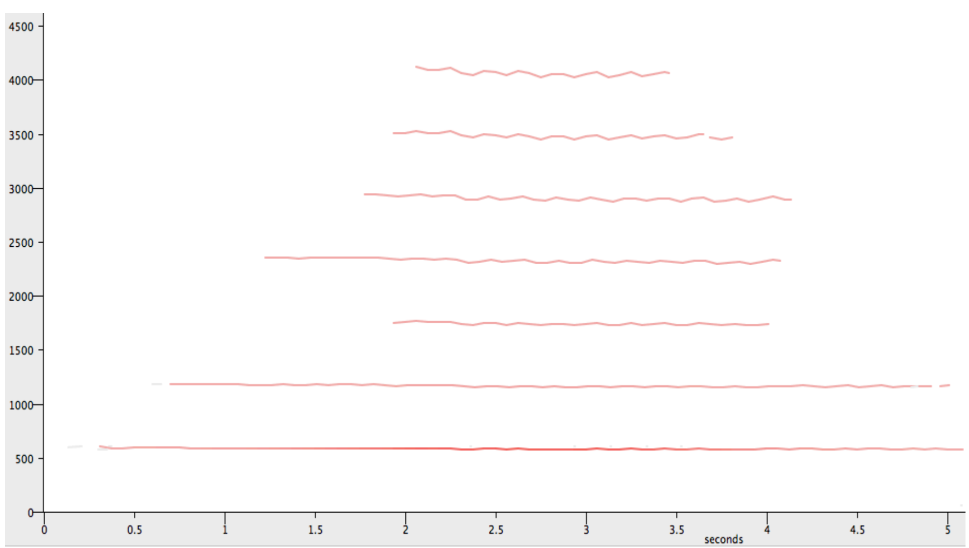

Figure 29: X-3 Channel 3 Tail Partials--Isolated Vocal Timbre.

Once X-3's tail vocal timbre has been isolated, a female mezzo-soprano voice moves to the foreground. Its "realness" is striking. All of Chowning research is evident in this example. At the attack level, a slight pitch variation is detected and even though a periodic vibrato is apparent, it is un-mechanized through the use of a semi-random vibrato.

This vocal sound here was generated by multi-carrier FM synthesis, on which algorithm 6 of TX81Z is based, which is referred to in the ChucK manual and the STK toolkit as FMVoices.[117] Again, this is an extension of the FM class. Barry Vercoe also uses this algorithm in CSound as *fmvoice* opcode.[118] Chowning describes how the voice-generating algorithm can be extended:

> (1) Summing the outputs of two or more copies of the basic algorithm in parallel, (2) one carrier oscillator and tow or more modulating oscillators in parallel, (3) one carrier oscillator and two or more modulating oscillators in series, and (4) two or more oscillators and one modulating oscillator … It is the last of these that is particularly appropriate to voice synthesis or indeed any tones that have prominent resonances.[119]

Applying Chowning's first procedure in Pd is illustrated in Figure 30. All three FM generators run in parallel and are added to a sum bus similarly to the additive synthesis principle earlier.

---

[117] Cook and Wang, "The Chuck Manual," 156.
[118] Vercoe, 747-748.
[119] Chowning, "Frequency Modulation Synthesis of the Singing Voice," 58.

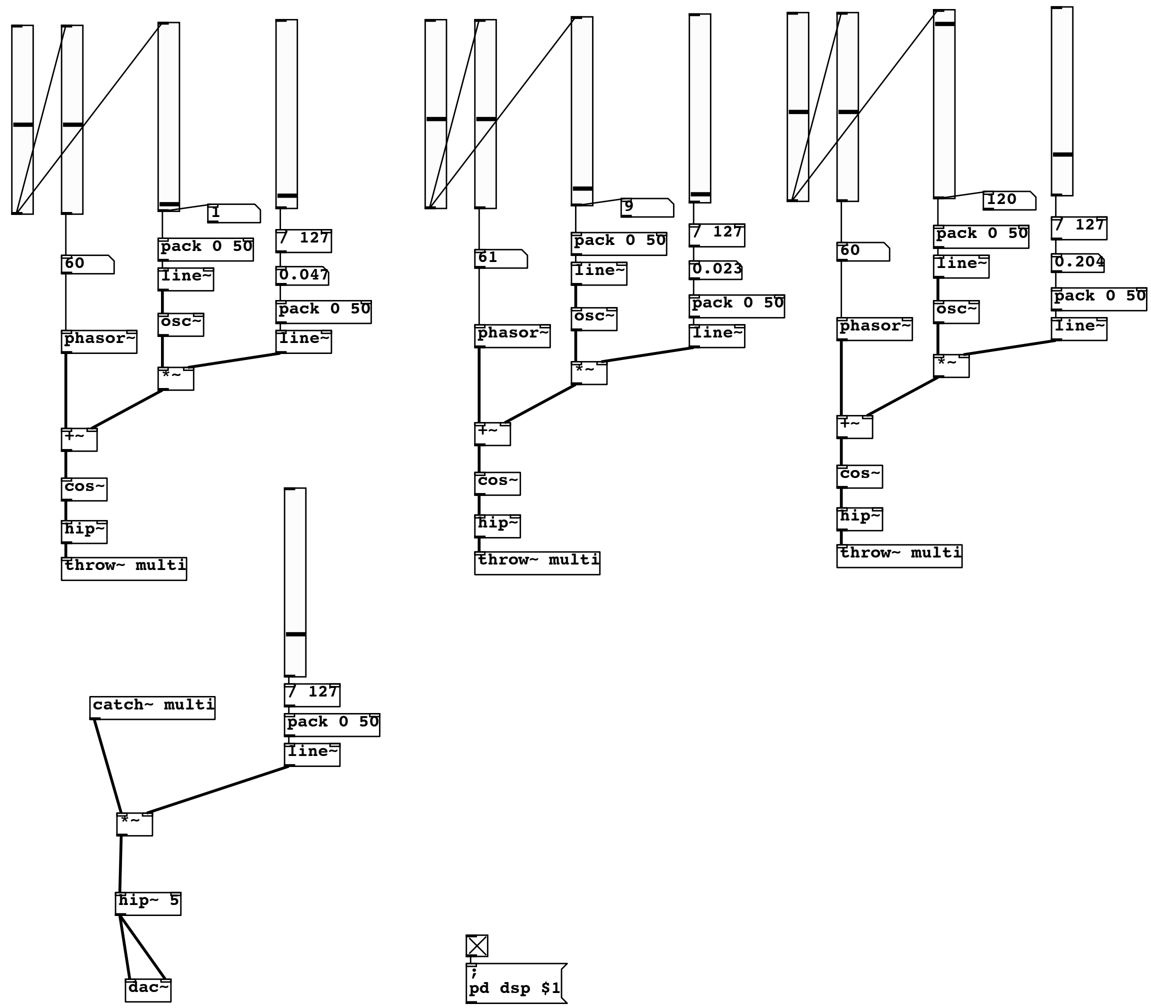

Figure 30: Procedure 1.

X-3 is different from the previous occurrences of X in that a metallic sound itself is not transformed to a vocal sound. The head (metallic sound) is separately generated from the tail (vocal sound). In a spatial multi-channel environment, the four channels are mixed to create an illusion of coherence. In effect, X-3 nonetheless can map onto X.

X-5

On initial observation, X-5 looks similar to X-3 or X-1. The length of X-5 is about 3.5 seconds and its exact position within A is shown in Table 10 of Appendix A. Figure 31 shows how the waveform of X-5 looks with all channels mixed down to mono.[120]

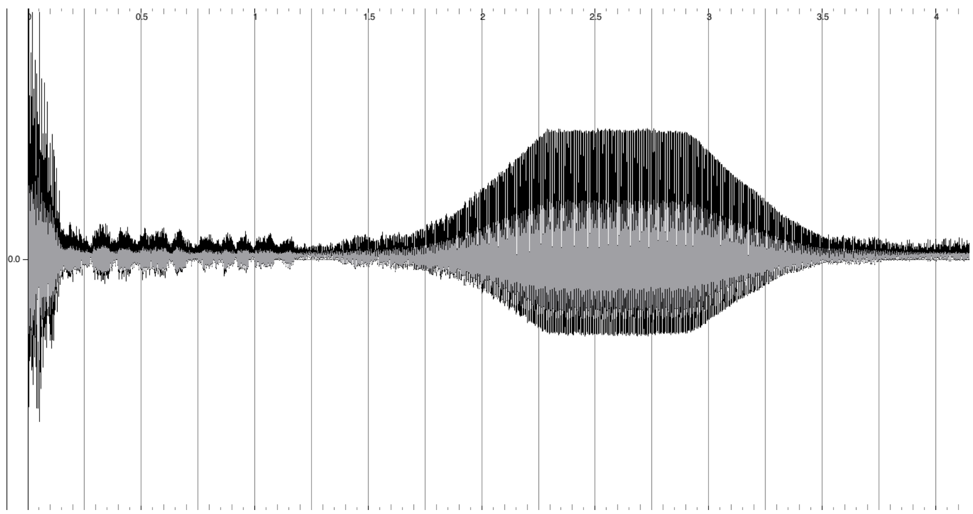

Figure 31: X-5 Waveform (mono).

From this perspective (intensity envelope – y-axis) X-5 also maps onto the proposed X object (Figure 2). The head section, lasting about 1.25 seconds, displays the sharp attack and quick decay phase previously observed with the head sections of X-1, X-3 and the "heads" of the preliminary opening gesture and represents the metallic bell tone that Chowning describes in his 1973 article "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation."[121] The tail section has an even attack (1.25 to 2.3 seconds on the x-axis), a leveled

[120] Again, a mono mix down was chosen to illustrate a unified perception.
[121] Chowning, "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation," 533.

sustain (2.3 to 2.9 seconds) and a decay section (2.9 to 3.75 seconds) that mirrors the attack phase in retrograde.

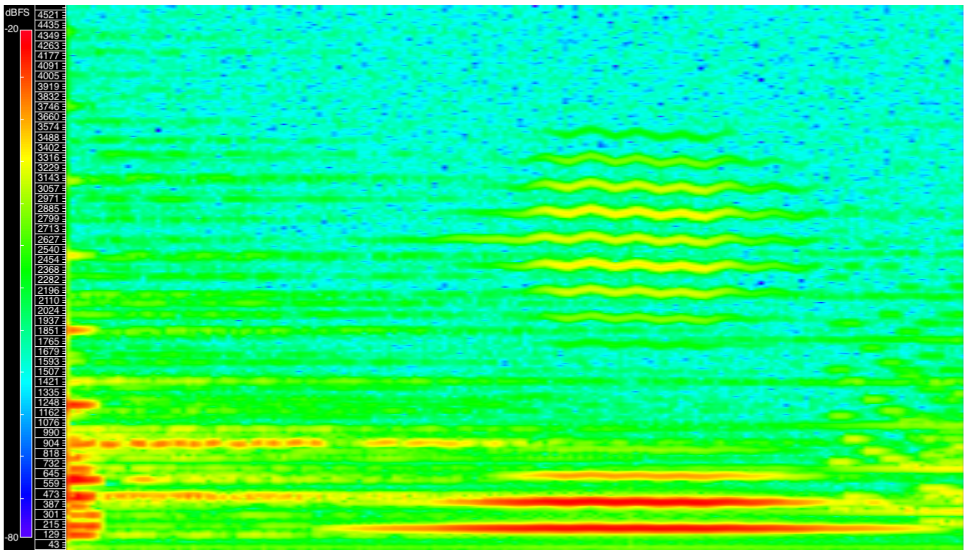

Figure 32: X-5 Spectrogram.

The spectrogram of X-5 confirms the findings about the head and tail sections of X-5 (Figure 32). The head section shows high intensity levels, in red and orange, at the attack level that highlight the inharmonic nature of the metallic or bell sound. Through reverberation the partials are extended into the sustain phase of the tail section. The tail section fades slowly in an out. The low fundamental partial is at the bottom and carries the highest intensity levels at around 200 Hz. This partial also lasts the longest. The two partials above are harmonic in character and are equally spaced at about 400 and 600 Hz respectively. These three partials also display the application of vibrato which is exemplified in their wavy attributes.

As in X-1 and X-2 an elliptical orb can be observed right above the audible 3 partials ranging from ca. 2000 Hz to ca. 3500 Hz. These are equally spaced as well, indicating their harmonic character. The most intense partial is situated at ca. 2700 Hz, meaning it is the carrier frequency. The remaining partials of this group are the resulting side bands. This procedure clearly indicates the use of a multi carrier FM synthesis algorithm configuration, similar to algorithm 6 (TX81Z) and that described by Chowning in his article “Frequency Modulation of the Singing Voice,” with the exception that now the voice sounds like a tenor.[122] Curtis Roads explains this algorithm as *Multiple-Carrier FM* or *MC FM*.[123] Dodge and Jerse describe the algorithm as *Double-Carrier FM Instruments*.[124] Most recently, the same procedure is also described in Julius O. Smith’s article “FM Voice,” in which he states that the procedure “can be viewed as *compressed modeling of spectral formants*.”[125] Smith further elaborates:

> A basic FM patch, consisting of two sinusoidal oscillators (a “modulator” and a “carrier’ oscillator”) can synthesize a useful approximation to a formant group in a harmonic line spectrum. In this application, the carrier frequency is set near the formant center frequency, and the modulating frequency is set to the desired pitch (e.g., of a sung voice).[126] The modulation index is set to give the desired bandwidth for the formant group. For the singing voice, three or more formant groups yield a sung vowel sound. Thus, a sung vowel can be synthesized using only six sinusoidal oscillators using FM. In straight additive synthesis, a bound on the number of

---

[122] Chowning, "Frequency Modulation Synthesis of the Singing Voice," 57-63. Chowning only describes the synthesis of a soprano singing voice, but the same “reality” effect can also be achieved for lower voices.
[123] Roads, *The Computer Music Tutorial*, 236-239.
[124] Dodge, 128-135.
[125] Julius O. Smith III, "Fm Voice", Center for Computer Research in Music and Acoustics, Stanford University https://ccrma.stanford.edu/~jos/sasp/FM_Voice.html (accessed 11.22.2009).
[126] Chowning, "Frequency Modulation Synthesis of the Singing Voice," 60.

> oscillators needed is given by the upper band-limit divided by the fundamental frequency, which could be, for a strongly projecting deep male voice, on the order of kHz divided by 100 Hz, or 200 oscillators.[127]

Chowning provides additional insight into the FM synthesis of vocal sounds in section 6.4 in his article *Frequency Modulation of the Singing Voice*:

> For the FM model of sung soprano tones three formants are considered. One oscillator can be used to modulate the three carrier oscillators with a separate index scaling for each, or three parallel FM pairs can be used. The frequency of the modulating oscillator(s) is always set to the frequency of the pitch of the tone, f0, while the frequencies of the carrier oscillators are set to those harmonic frequencies closest to the appropriate formant frequencies … the FM equations are computed from the basic musical descriptors of overall amplitude and fundamental pitch and from a set of tables that form the data base for the terms at selected pitches through the soprano range.[128]

Figure 33 shows the exact partials contained in X-5. Here, the vibrato fluctuations are clearly outlined in the elliptical orb (1.5 to 3.5 seconds on the x-axis).

---

[127] Smith III, "Fm Voice."
[128] Chowning, "Frequency Modulation Synthesis of the Singing Voice," 59.

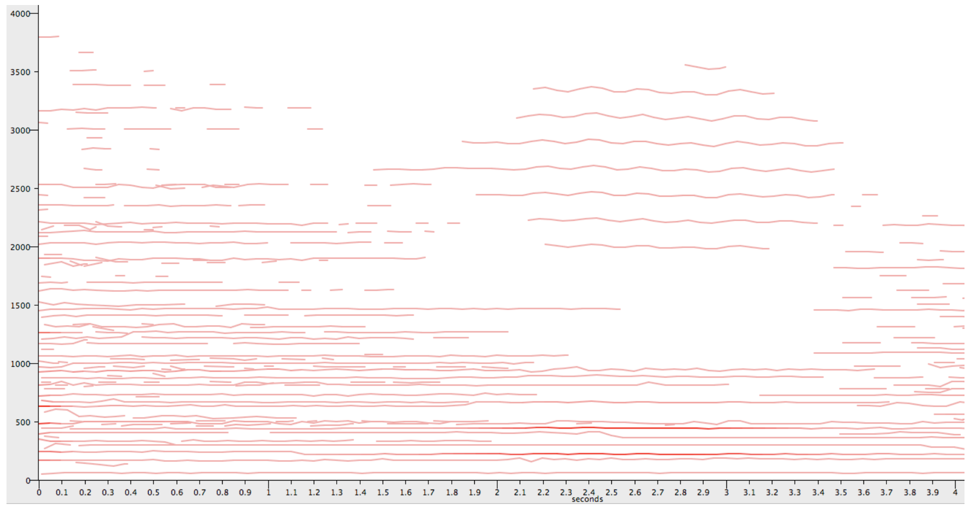

Figure 33: X-5 Partials.

A further examination of object X-5 reveals that Chowning uses the same procedure, in regards to channel distribution, as with object X-3. Channels 1 and 2 contain the head (metallic sound) of X-5 (Figure 34).

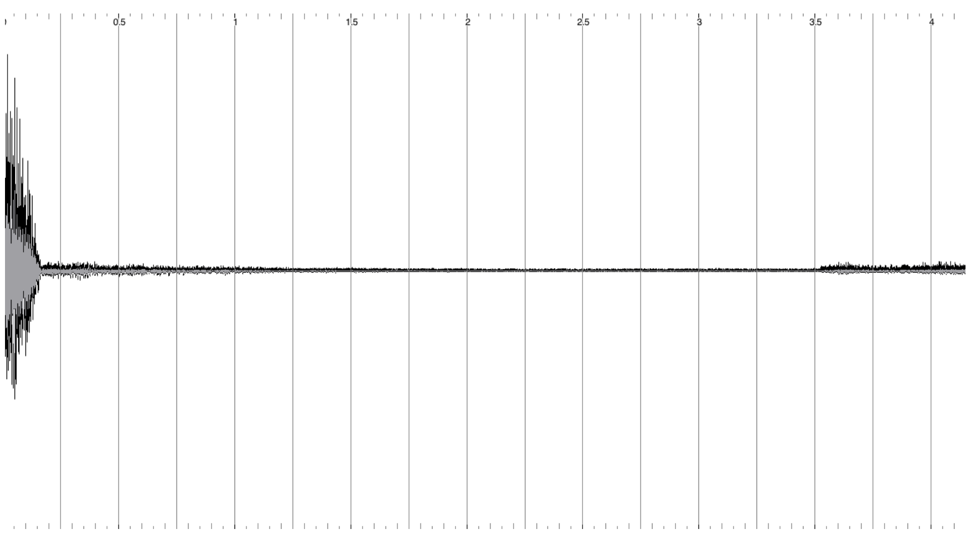

Figure 34: X-5 Channel 2 (Head).

Channels 3 and 4 contain the tail of X-5, the vocal sound (Figure 35).

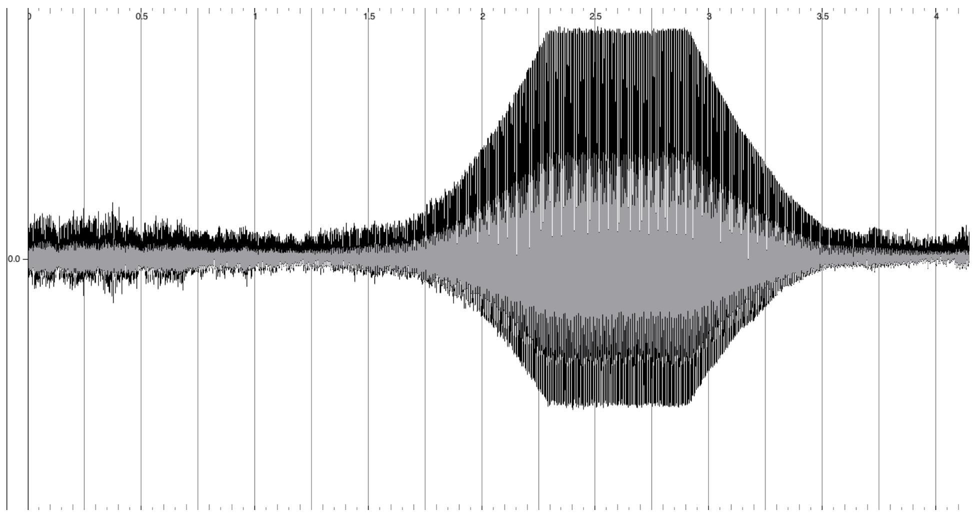

Figure 35: X-5 Channel 3 (Tail).

However, channel 1 is at a lower intensity level than channel 2 and channel 4 is at a lower intensity level than channel 3. In both instances, it seems as if the actual sounds originate in channels 2 and 3. As with X-3, the tail and head sections a generated separately and do not actually transform from one sound into another. Through clever mixing procedures of the 4 channels, X-5 fuses together as a clear representation of X. A patch that illustrates how to generate a vocal-like sound has already been illustrated (Figure 30).[129]

[129] Another issue that does not necessarily pertain to this analysis is how to connect the generated sounds with a MIDI keyboard. APPENDIX B - *PD* PATCHES, Figure 44 illustrates how I was able to connect my personal MIDI keyboard to an FM synth with ADSR control routed over sliders. The patch works in conjunction with Figure 42.

CONCLUSION

The Pythagoras legend tells a great story of discovery. The discovery that tones and notes can be related to each other by size. His profound insight led to the ratios of the harmonic series and has influenced the course of music for millennia. As I learned from this study the harmonic series of a metal object is inharmonic. Nonetheless, I find it fascinating that Pythagoras was able to describe the basic precepts of the nature of timbre. It is the spirit of Pythagoras that drives computer musicians, enthusiasts and hobbyists alike to the frontier of new explorations of timbres.

In this study, I evaluated John Chowning's composition *Phoné*. I chose *Phoné* for this study, because it is a fixed media piece, which means that its attributes are limited to a recorded realization. I defined an unknown parameter suitably called X. This parameter functioned as a placeholder of what I would find. I limited the parameter to a very small aspect of the composition, because computer music can contain parameters *ad inifinitum*. A computer music composer can spend many years to assemble a new composition due to this challenge.

After that, I contextualized these parameters and their setting within the composition, by examining their dynamic, pitch, harmonic and spectral content via FFT analysis in *SPEAR* and *Sonic Visualizer*. Then, I connected these parameters with articles the composer, teachers and researchers had written about the subject manner. Subsequently, I was able to build a few patches in *Pd* that helped me understand these concepts through experimentation and listening.

Furthermore, I chose *Pd* as an *auralization* tool, because of its intuitive nature and ease of use. In *Pd* complex mathematical expressions are calculated with the use of just a few objects and can be programmed on the fly. I also examined other music programming environments such as *Csound* and *ChucK* and found that these environments require a great deal of programming knowledge and prowess. In some instances, it would have been easier for me to re-create certain patches in these environments since they are based on Max Matthews's famous Music I-XI program and Chowning used these environments himself. These tools are great and any of the findings presented here can also be programmed in these environments successfully. However, from a conceptual point of view I believe it is *Pd*'s simplicity, infinite expandability and community that make it a very elegant instrument in the music theorist's toolbox.[130]

With these tools and the proposed procedure, I was able to find and explain attributes of *Phoné* that would otherwise not have been possible. Many of these findings were surprising, yet satisfying and some were unexpected. However, I was able to associate all of my findings with the X parameter.

---

[130] John Chowning's most recent composition *Voices* was realized using Max/MSP, the commercial cousin of *Pd*. As previously stated, Chowning advises to use Max family programming environments in connection with the JOS index.

Phoné is a 4-channel composition. 4-channel compositions have been part of the electroacoustic music composer's tool set since *Gesang der Jünglinge* by Karlheinz Stockhausen (1956).[131] Chowning draws on this tradition and transfers it to the computer.[132] I found that Chowning's techniques involve spatialization of sound sources that crystallize through his dispersal of timbral material among different channel groups. In some cases, *Phoné*'s channel separation was used to unify two different synthesis techniques. Chowning juxtaposed FM synthesis technique for metal sounds with the synthesis of vocal sounds. This was manifested in X.

I also found that Chowning juxtaposed these two timbres with the use of the multi carrier FM synthesis algorithm. Chowning transforms metal timbre to voice timbre, or at least creates illusions of such transformations. At times, the vocal sounds can have more of a metal quality and at other times, a metal timbre can project vocal properties. In order to achieve these transformations, Chowning also uses more basic synthesis techniques such as different configurations of an ADSR generator and reverb. These techniques were manifested in X as well.

These techniques are projections of Chowning's aesthetic to create music that is unique to the computer. Chowning's idea on *perceptual fusion* manifested itself in the spatialization and choice of pitch material. The pitch material was not chosen at random by a patch or a row, but rather it was derived from the partials

---

[131] Chadabe, 39.
[132] John M. Chowning, "The Simulation of Moving Sound Sources," *Computer Music Journal* 1, no. 3 (1977).

of the harmonic series of discrete pitches. If the partials occurred in successive order, they were serialized and would end with a simultaneity that could be a chord. The generated chord often resulted in a timbre through perceptual fusion.

This technique and aesthetic was closely aligned with Risset's composition *Mutations*. Risset made use of the same technique of linking timbre to pitch space. Risset shared his view of a *Gestalt* phenomenon with Pierre Schaeffer. In fact conceptually, I found that this compositional theory could be traced to Schoenberg. The attributes of X are rooted within this compositional philosophy.

However, to merely view the techniques of computer music composers as music theoretical entities in themselves does not guarantee any meaningful analytical results. To be able to hear the technical and aesthetical concepts and to be able to recreate them will contribute greatly to the analysis of electroacoustic music, computer music and music at large.

APPENDIX A - EVENT CHART

X – Only Event List

ID	Item	Duration	Start Time	End Time	Section	Attributes		X
1	X	00:29.950	00:15.387	00:45.337	A		Single	1
2	X	00:29.275	00:58.032	01:27.307	A		Single	2
3	X	00:05.086	01:35.161	01:40.247	A		Single	3
4	X	00:23.639	01:46.493	02:10.133	A		Single	4
5	X	00:03.513	02:12.414	02:15.927	A		Single	5
6	X	01:33.535	03:32.405	05:05.940	A		Single	6
7	X (Channel-3+4)	00:20.942	05:15.512	05:36.454	B	Imitative - long	Group-1	7.1
8	X (Channel-1)	00:12.906	05:23.549	05:36.454	B	Imitative - Medium	Group-1	7.2
9	X (Channel-2)	00:02.747	05:28.452	05:31.200	B	Imitative - Shortest	Group-1	7.3
10	X (Channel-2)	00:03.662	05:36.964	05:40.627	B	Imitative - Short	Group-1	7.4
11	X (Channel-3)	00:16.568	06:27.898	06:44.466	B	Imitative - Similar Length	Group-2	8.1
12	X (Channel-4)	00:07.606	06:35.965	06:43.571	B	Imitative - Delayed -Similar Length	Group-2	8.2
13	X (Channel-1)	00:26.313	06:36.163	07:02.475	B	Imitative - Delayed -Similar Length	Group-2	8.3
14	X (Channel-2)	00:26.587	06:36.209	07:02.795	B	Imitative - Delayed -Similar Length	Group-2	8.4
15	X (Channel-1)	00:29.299	08:03.477	08:32.776	C	Synchronized	Group-3	9.1
16	X (Channel-4)	00:27.053	08:03.477	08:30.530	C	Synchronized	Group-3	9.2
17	X (Channel-3)	00:27.879	08:03.477	08:31.356	C	Synchronized	Group-3	9.3
18	X (Channel-2)	00:41.003	08:03.477	08:44.479	C	Synchronized	Group-3	9.4
19	X	00:33.066	09:03.477	09:36.542	C		Single	10
20	X	00:35.086	10:42.582	11:17.669	C		Single	11
21	X	00:04.423	11:19.472	11:23.895	D		Single	12
22	X	00:48.116	12:11.523	12:59.639	D		Single	13

Table 10: X Occurrences

APPENDIX B - *PD* PATCHES

An Initialization Function

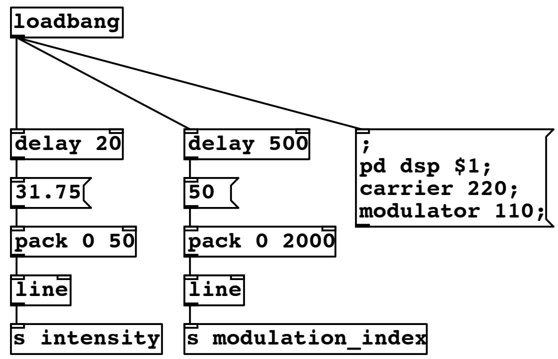

Figure 36: Contents of the `pd init` object from Figure 1.

Figure 36 shows the contents of the. This is a utility function that loads a few preset parameters to instantly start the patch from Figure 1. A `bang` is send when the patch is opened. This `bang` turns on the digital signal processor (`pd dsp 1`) and sends 220Hz to the carrier sinusoid and 110Hz to the modulation sinusoid. After 20 milliseconds (`delay 20`) another `bang` is send to the volume slider (`intensity`) that smoothly ramps up the in 50 milliseconds (`pack 0 50`, `line`) volume in 50 milliseconds to a specified value. This procedure avoids a ‘click’ sound at start-up, since an attack that immediately jumps to maximum capacity results in a percussive sound. The same procedure is used for the modulation index (`delay 500`), in order to demonstrate a sweep through the harmonic series (`pack 0 2000, line`) within the first 2.5 seconds of opening the patch.

Sample Player

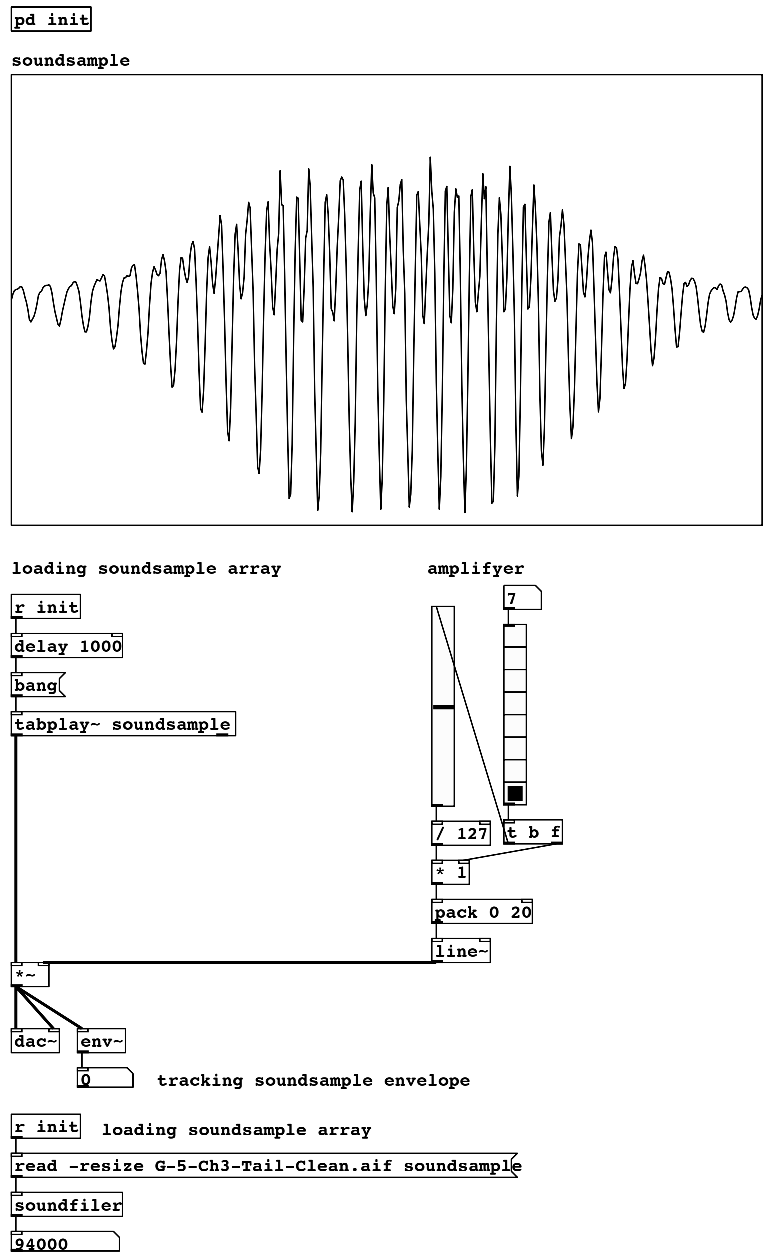

Figure 37: Simple Sample Player.

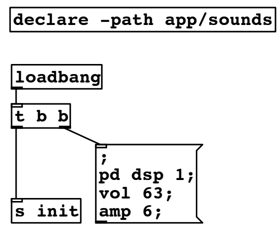

Figure 38: Initializing Function for previous patch.

Vibrato - Basic FM Synthesis

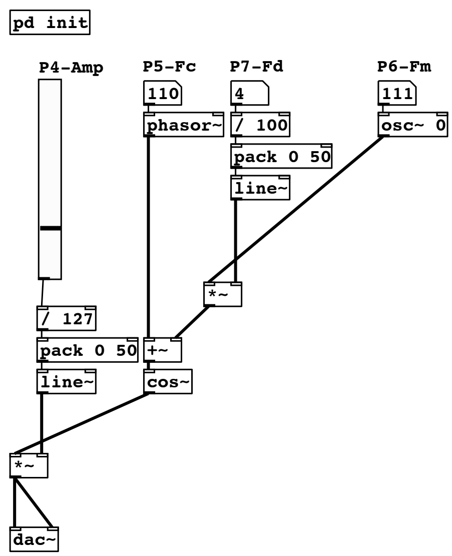

Figure 39: Basic FM Synthesis--in Pd after Chowning's 1973 article.

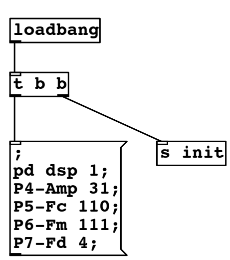

Figure 40: Initialization Function for previous patch.

FM Synthesis Spectral Envelope

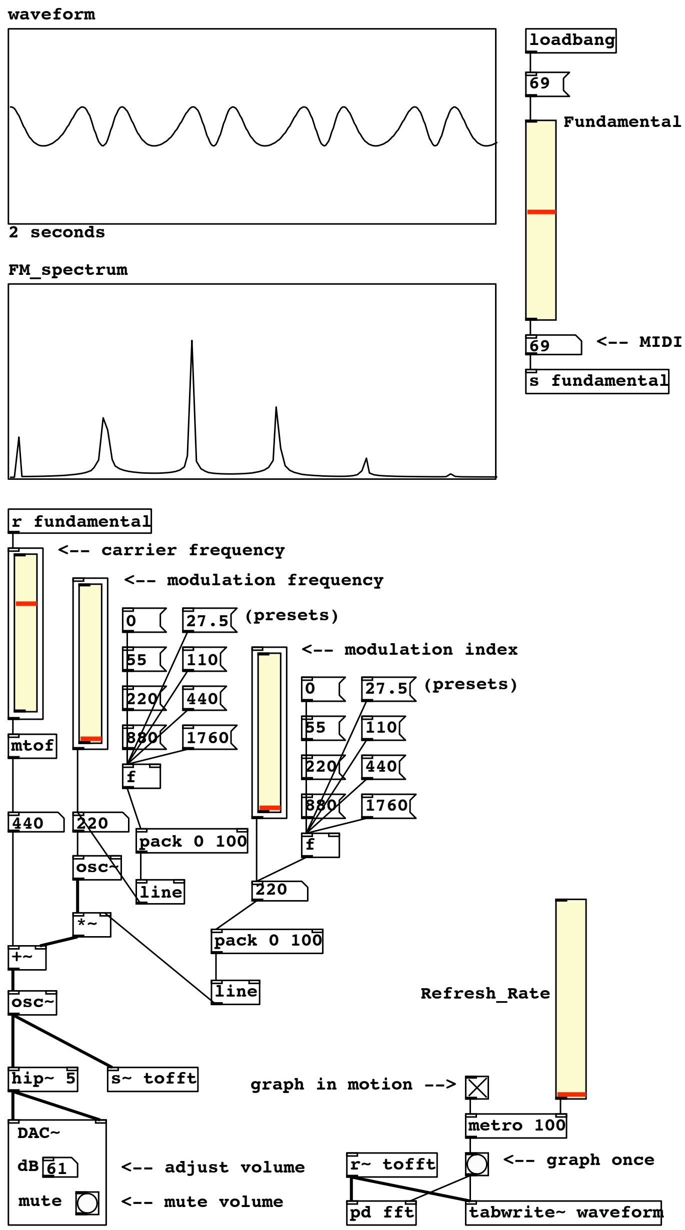

Figure 41: Plotting FM Synthesis Spectrum in Pd.

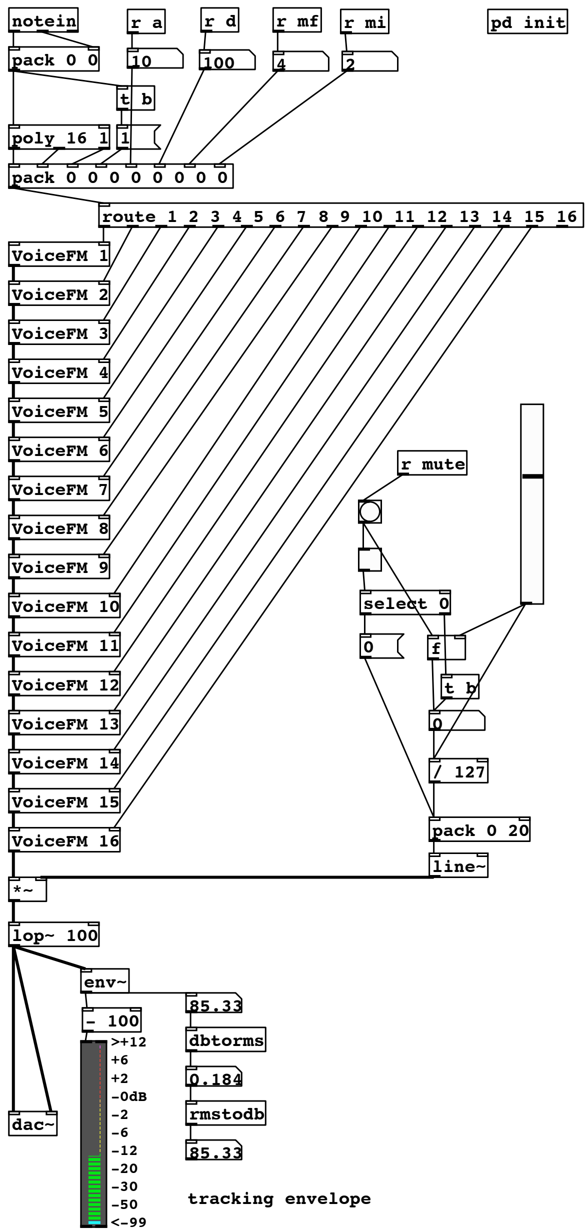

Figure 42: Creating Polyphony in Pd--Voices here are FM.

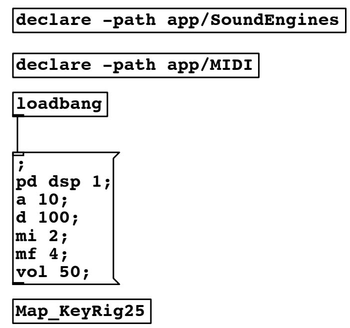

Figure 43: Initialization Function of Polyphony patch.

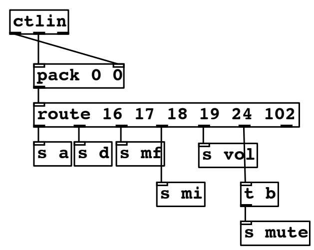

Figure 44: M-Audio KeyRig 25 MIDI mapping.

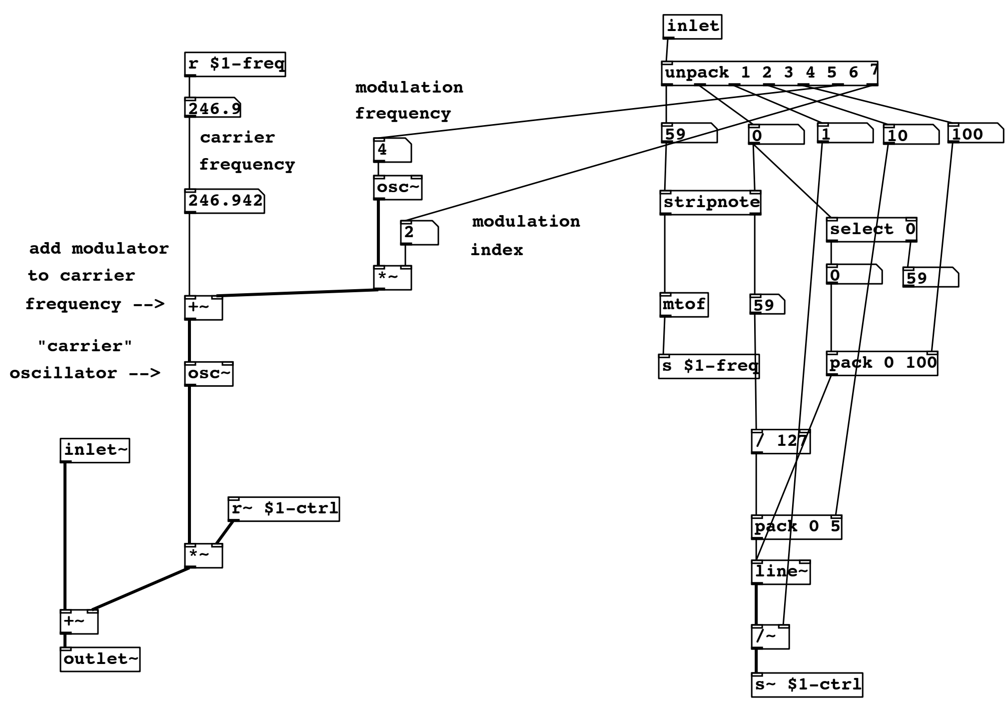

Figure 45: Individual Voice as used in Polyphony patch.

BIBLIOGRAPHY

Battier, Marc, Pierre Boulez, John Chowning, Hugues Dufourt, Brian Ferneyhough, Tod Machover, Benedict Mailliard, Marvin Minsky, and Iannis Xennakis, eds. *Le Compositeur Et L'ordinateur*. Paris: IRCAM en association avec l'Ensemble InterContemporain, 1982.

Boethius, Anicius Manlius Severinus, "De Institutione Musica, Liber 1", Indiana University http://www.chmtl.indiana.edu/tml/6th-8th/BOEDIM1_TEXT.html (accessed 03.08.2010).

Bower, Calvin M. "The Transmission of Ancient Music Theory into the Middle Ages " In *The Cambridge History of Western Music Theory*, edited by Thomas Christensen 136-167. New York: Cambridge University Press, 2007.

Bristow, David, and John Chowning. *Fm Theory and Applications*. Tokyo: Yamaha Music Foundation, 1986.

Chadabe, Joel. *Electronic Sound: The Past and Promise of Electronic Music*. Upper Saddle River, New Jersy: Prentice-Hall, 1997.

Chowning, John M. "DVD Program Notes." *Computer Music Journal* 32, no. 4 (2008): 106-118.

________. "The Synthesis of Complex Audio Spectra by Means of Frequency Modulation." *Journal of the Audio Engineering Society* 21, no. 7 (1973): 526-534.

________. "The Simulation of Moving Sound Sources." *Computer Music Journal* 1, no. 3 (1977): 48-52.

________. "Frequency Modulation Synthesis of the Singing Voice." In *Current Directions in Computer Music*, edited by Max V. Mathews and John R. Pierce, 57-63. Cambridge: MIT Press, 1989.

________. "Digital Sound Synthesis, Acoustics and Perception: A Rich Intersection." In *COST G-6 Conference on Digital Audio Effects*, 6. Verona, Italy, 2000.

________. "Fifty Years of Computer Music: Ideas of the Past Speak to the Future." In *Computer Music Modeling and Retrieval. Sense of Sounds: 4th International Symposium, Cmmr 2007, Copenhagen, Denmark, August 27-31, 2007. Revised Papers*, 1-10: Springer-Verlag, 2008.

Cook, Perry R., ed. *Music, Cognition, and Computerized Sound : An Introduction to Psychoacoustics*. Cambridge, Mass.: MITPress, 1999.

Cook, Perry R., and Ge Wang. *The Chuck Manual*. Princeton: Princeton University, 2007. http://chuck.cs.princeton.edu/release/files/chuck_manual.pdf.

Di Scipio, Agostino. "Jean-Claude Risset: Contours (1982)." In *Electroacoustic Music*, edited by Thomas Licata, 242. Westport, Connecticut: Greenwood Press, 2002.

Dodge, Charles and Thomas A. Jerse. *Computer Music*. 2nd ed. New York: Schirmer Books, 1997.

Fennelly, Brian. "A Descriptive Language for the Analysis of Electronic Music." *Perspectives of New Music* 6, no. 1 (1967): 79-95.

Fineberg, Joshua. "Guide to the Basic Concepts and Techniques of Spectral Music." *Contemporary Music Review* 19, no. 2 (2000): 81-113.

Gayou, Évelyne, ed. *John Chowning*. Vol. 7, Portraits Polychrome. Paris: Institut national de l'audiovisuel, 2007.

Hamburger Jahrbuch Für Musikwissenschaft. Vol. 11. New York: P. Lang, 1991.

Harvey, Jonathan. "'Mortuos Plango, Vivos, Voco': A Realization at Ircam." *Computer Music Journal* 5, no. 4 (1981): 22-24.

Jones, David Evan. "Compositional Control of Phonetic/Nonphonetic Perception." *Perspectives of New Music* 25, no. 1/2 (1987): 138-155.

Klingbeil, Michael Kateley. "Spectral Analysis, Editing, and Resynthesis: Methods and Applications." D.M.A. diss., Columbia University, 2009.

Kochan, Stephen G. *Programming in Objective-C 2.0*. Upper Saddle River, NJ: Addison-Wesley, 2009.

Kohl, Jerome. "Stockhausen: Telemusik (1966)." In *Electroacoustic Music*, edited by Thomas Licata, 242. Westport, Connecticut: Greenwood Press, 2002.

Landy, Leigh. *Understanding the Art of Sound Organization*. Cambridge: MIT Press, 2007.

Larsen, Steve. "The Problem of Prolongation In 'Tonal' Music: Terminology, Perception, and Expressive Meaning." *Journal of Music Theory* 41, no. 1 (1997): 101-136.

Licata, Thomas, ed. *Electroacoustic Music*. Westport, Connecticut: Greenwood Press, 2002.

Manning, Peter. *Electronic and Computer Music*. New York: Oxford University Press, 2004.

Mathews, Max V., and John R. Pierce, eds. *Current Directions in Computer Music Research*, System Development Foundation Benchmark Series. Cambridge, Massachusetts: The MIT Press, 1989.

Moore, Richard F. "Dreams of Computer Music: Then and Now." *Computer Music Journal* 20, no. 1 (1996): 25-41.

Morgan, Robert P. *Twentieth-Century Music: A History of Musical Style in Modern Europe and America*. New York: W. W. Norton, 1991.

Moscovich, Viviana. "French Spectral Music: An Introduction." *Tempo*, no. 200 (1997): 21-27.

Puckette, Miller. "Max at Seventeen." *Computer Music Journal* 26, no. 4 (2002): 31-43.

________. *The Theory and Technique of Electronic Music*. Hackensack, NJ: World Scientific, 2007.

Risset, Jean Claude. "Computer Music Experiments 1964 - ... [sic]." *Computer Music Journal* 9, no. 1 (1985): 11-18.

Risset, Jean-Claude. "The Perception of Musical Sound." http://www.utexas.edu/cola/insts/france-ut/_files/pdf/resources/risset.pdf.

Roads, Curtis. *The Computer Music Tutorial*. Cambridge, Massachusetts: The MIT Press, 1996.

________. *Microsound*. Cambridge: MIT Press, 2001.

________, ed. *Composers and the Computer*. Los Altos, Calif.: W. Kaufmann, 1985.

Rose, François. "Introduction to the Pitch Organization of French Spectral Music." *Perspectives of New Music* 34, no. 2 (1996): 6-39.

Salluce, Antonio, "John Chowning", Institut national de l'audiovisuel http://www.ina-entreprise.com/sites/ina/medias/upload/grm/portraits-polychromes/extraits/chowning/index.html (accessed 03.07.2010).

Schaeffer, Pierre. *Traité Des Objets Musicaux: Essai Interdisciplines*. 2nd ed. Paris: Seuil, 1977.

Schiano, Michael J., "Grundgestalt", Grove Music Online. Oxford Music Online. http://www.oxfordmusiconline.com/subscriber/article/grove/music/11868 (accessed 10.06.2009).

Schoenberg, Arnold. "Linear Counterpoint." In *Style and Idea*, edited by Leonard Stein, 559. Berkeley: University of California Press, 1984.

________. *The Musical Idea and the Logic, Technique and Art of Its Presentation*. Translated by Patricia Carpenter and Severine Neff, Edited by Patricia Carpenter and Severine Neff. Bloomington, Ind.: Indiana University Press, 2006.

Simoni, Mary, ed. *Analytical Methods of Electroacoustic Music*. New York: Routledge, 2006.

Smith III, Julius O., "Global Jos Index," Center for Computer Research in Music and Acoustics, Stanford University https://ccrma.stanford.edu/~jos/GlobalJOSIndex.html (accessed 03.07.2010).

________, "Fm Voice", Center for Computer Research in Music and Acoustics, Stanford University https://ccrma.stanford.edu/~jos/sasp/FM_Voice.html (accessed 11.22.2009).

________, "Sinusoidal Frequency Modulation (Fm)", Center for Computer Research in Music and Acoustics, Stanford University https://ccrma.stanford.edu/~jos/mdft/Sinusoidal_Frequency_Modulation_FM.html (accessed 03.07.2010).

Sundberg, Johann. "Synthesis of Singing." *Swedish Journal of Musicology* 60, no. 1 (1978): 107-112.

Vercoe, Barry. *The Canonical Csound Reference Manual: Version 5.10*: MIT Media Lab.

Wishart, Trevor. *On Sonic Art*, Edited by Simon Emmerson. Amsterdam: Harwood Academic Publishers, 1996.