Mathews' Diagram and Euclid's Line —Fifty Years Ago—

John Chowning
CCRMA, Stanford University
jc@ccrma.stanford.edu

ABSTRACT

Making the science and technology of computer music comprehensible to musicians and composers who had little or no background therein was a part of Max Mathews' genius. In this presentation I will show how a simple diagram led to the essential understanding of Claude Shannon's sampling theorem, which in turn opened up a conceptual path to composing music for loudspeakers that had nothing to do with wires, cables and electronic devices, but led to learning how to program a computer—to write code. The change from device-determined output (analog) to program-determined output (digital) was a major change in paradigm that led to my realization of an integral sound spatialization system that would have been impossible for me to achieve in any other medium. Along the way, the discovery of FM Synthesis provided not only a means of creating diverse spectra but coupled with a ratio from Euclid's Elements produced an unusual and productive connection between spectral space and pitch space and a path that leads …?

1. INTRODUCTION

Claude Shannon's 1948 paper, "A Mathematical Theory of Communication" [1] is the hard-edged theory that underlies the flow of information in today's complex digital world of computers, large and small, tablets, mobile phones, pads and pods—capable of "sensing" sound, image, touch, location —all complex machines, the complete understanding of which is beyond the capacity to know of any single human being. It is a summation of Shannon's own work and that of his colleagues and predecessors. The timing was propitious as the first stored-program computers were just being developed. The paper includes the first use of the word bits.[1] And theorem 13, the sampling theorem, is critical to the connection between continuous and discrete signals. In his article, "The origins of the sampling theorem," H.D. Luke traces the rich history of the sampling theorem that extends back to 1848 [2].

[1] "If the base 2 is used the resulting units may be called binary digits, or more briefly bits, a word suggested by J. W. Tukey" [1].

Shannon's paper is the first reference in Max Mathews' famous 1963 article "The Digital Computer as a Musical Instrument" [3], because the sampling theorem is the foundation on which Mathews based much of his early work. His research included speech, hearing and computer music where the loudspeaker is the ultimate sound source. Mathew's diagrammatic representation of the sampling theorem opened the door to my understanding of what was otherwise incomprehensible because of my own "nothing-but-music" background.

Euclid's line, to which I refer in the title, is its division into extreme and mean ratio now commonly known as the Golden ratio. This ratio became of interest to me after composing Turenas (1972), in which I made extensive use of both harmonic and inharmonic spectra. I looked for other irrational numbers to produce inharmonic spectra and found that the Golden ratio had particularly interesting properties in this application.

2. MATHEWS' DIAGRAM

My interest in music composed for loudspeakers stemmed from a few musical experiences that had a profound effect on the way I thought about music. From 1959 until 1962 I studied in Paris where contemporary music was notably present. Some concerts included electroacoustic music— the Domaine Musicale concerts at the *Théâtre de l'Odéon* and the Groupe de recherches musicales (GRM) presented concerts at the French Radio that were exclusively electroacoustic. Some of the music, composed for 4-channels was, quite literally, head turning. From my youth I had a fascination with cavernous spaces and echoes, their disorienting effect on otherwise familiar sounds and the spatial aspect of this music provoked a desire to compose for loudspeakers—imagined sounds in imagined spaces.

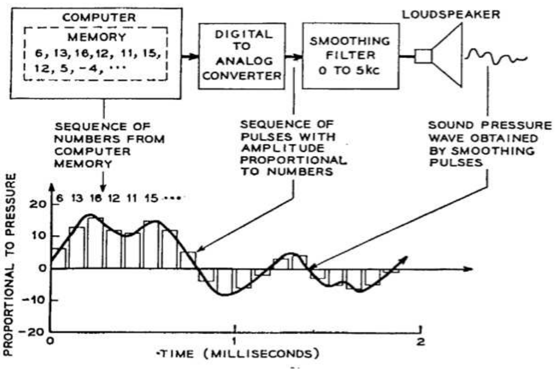

Figure 1. This is Mathews' schematic diagram of the sampling process from 1963 [3], at which time electro-acoustic music was exclusively in the analog domain.

But, I was well aware that the stringent technical requirements, knowledge and means to create music for loudspeakers in the 1960s were accessible to very few composers.

In 1964, because of a bit of serendipity, I was given Mathew's article. It was the first diagram, which caught my attention, see Figure 1. It presented a comprehensible face of the sampling theorem, which for me, and perhaps others, was suggestive and inspiring—even poetic in that it showed a path to electroacoustic music that bypassed what for me was technological clutter, a path that would allow the composition of "any perceivable sound" [3] bringing musical creation to the edge of my imagination.

2.1 Sampling's Simplicity

Immediately striking in the diagram is that there are but three devices and a computer, none of which have changed over time in their functional relationship, but all of which have changed over time in their cost, quality and precision—for the better!

Now, dear reader, imagine a 29 year old graduate student composer, fifteen years from his last math class, never having seen a computer, but with vivid imaginings, however vague and inchoate, of composing music in space. Imagine further, the conceptual breakthrough when with images in mind of electroacoustic music studios—filled with electronic equipment, cables, wires, multiple microphones, spinning loudspeakers and stern-looking engineers in white coats—I understood the implications of Mathew's Figure 1.

2.2 The Soft Complexity Behind the Samples

Already familiar with complex symbols as representation of sound, musicians seemed to be undaunted by learning to program a computer to do the same. Having read Mathew's article early in 1964 and the comprehensive article by James Tenney, "Sound-generation by means of a digital computer" [4], in April, I took a new course offered at Stanford University "Computer Programming for Non-Engineers." With the confidence that I could program a computer, I set about to learn acoustics and psychoacoustics, the latter highlighted in Mathew's article as an area of special importance to music perception.

Tutored by an undergraduate math major, tuba player, and incipient hacker, David Poole—my angel—by September 1964 (just 50 years ago!) we had generated our first sounds using Mathew's Music IV program. [2]

The Artificial Intelligence Laboratory provided me off hour computer time and a population of skilled researchers in fields ranging from linguistics to philosophy, speech, physics and, of course, computer science and electrical engineering, any one of whom could answer the many questions that I posed as I developed a sound spatialization program. I realized a quad system in 1968, after cajoling an EE student to build a 4-channel DAC with the promise of sounds swirling and swooping from everywhere, see Figure 2.

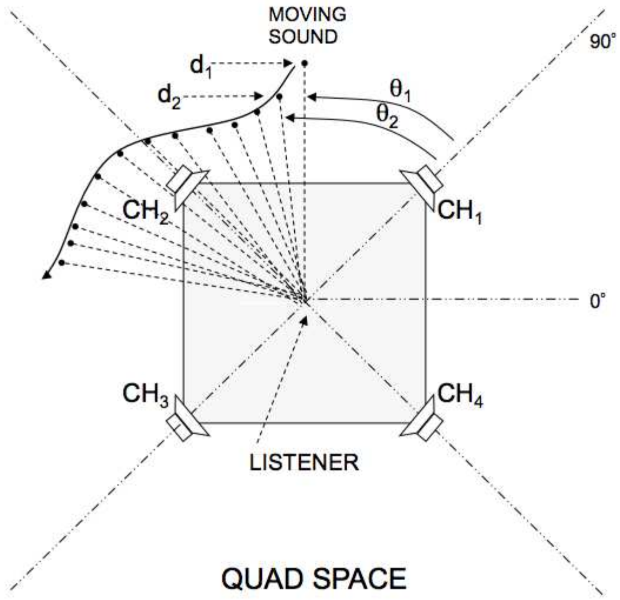

Figure 2. Finding a graphic solution: the distance, azimuth and velocity cues of a moving sound are captured by plotting points along the trajectory at a constant interval of time. Doppler shift is derived from the radial velocity. I used the Cartesian quadrants for naming the channels.

Completing the quad spatial system was a very important moment in my personal history and in the direction that the Computer Music Project—and eventually CCRMA—would take, for several reasons:

- While computers were not yet powerful enough to synthesize and process sound in real-time—hands-on and favoring immediate response—they would be some day (as we know very well with today's technology).
- Computer synthesis provided the composer direct control of the material of music, as a painter has with paint and canvas, allowing the accomplishment of two very different but complementary processes — joining the structure of the sound itself to the structure of musical form.
- I realized that those having motivation and perseverance, but no special competence in building electronic devices, were presented with a means to engage in a medium, and at a high level of abstraction, that was a defining musical advance in the 20th century—music composed for loudspeakers.

The discovery of FM Synthesis in 1967 was the result of searching for lively sounds that had some internal dynamism that made them easy to localize. Over the next few years I developed FM synthesis with Jean-Claude Risset's analysis-based synthesis of trumpet tones [5], providing a key insight. [3]

[2] The program was run on an IBM 7094, a 1301 disk drive, which was shared with a Digital Equipment Co. PDP-1, whose graphics display's x, y ladders provided DACs.

[3] Joined by Leland Smith, then in 1968 by J. "Andy" Moorer and then later by John Grey and Loren Rush, the research at the Computer Music Project flourished. The Center for Computer Research in Music and Acoustics (CCRMA) was founded in 1974.

After seven years of development and study, I had acquired the knowledge and built the tools to a sufficient level of sophistication to realize two compositions— Sabelithe (1971) and Turenas. An extensive account of this early work, "Turenas: the realization of a dream," was presented at the *Journées d'Informatique Musicale* in 2011 [6].

3. EUCLID'S LINE

Euclid defines what is now known as the Golden ratio in Elements, Book VI, Definition 3 [8].

> A straight line is said to have been cut in extreme and mean ratio when, as the whole line is to the greater segment, so is the greater to the less.

$$AB : AC = AC : CB \tag{1}$$

or

$$\begin{aligned} &= \frac{1+\sqrt{5}}{2} \\ &= 1.618033... \end{aligned} \tag{2}$$

The ratio in its algebraic form (equation 2) is one of the most studied of numbers, with many claims made over centuries in regard to its presence in nature, art, music, etc.—many are probably extravagant claims. The ratio is implicit in the formation of the pentagram and perhaps known to the Pythagoreans almost three centuries earlier. However, my interest in this ratio came from another point of view.

3.1 The Golden ratio and FM Spectra

In FM synthesis the distribution of the spectral (sideband) components are determined by the relationship between the carrier and the modulating frequencies. For inharmonic spectra in Turenas, I used a carrier frequency to modulating frequency ratio of 1:√2. Looking for other irrational numbers that satisfied the constraint that their fractional part not be small, as is, for example, π, I explored the sound and attributes of the Golden ratio. When the carrier and modulating frequencies are both different powers of φ, four of the resulting partials are

SIDEBAND (SB) FREQUENCIES FOR f_c=1000 * φ^0 and f_m=1000 * φ^1				
order	Lower SB			Upper SB
0	* Hz	1000	f_c	Hz
1	618.03	f_c-f_m	f_c+f_m	2618.03
2	2236.07	f_c-2f_m	f_c+2f_m	4236.07
3	3854.10	f_c-3f_m	f_c+3f_m	5854.10
4	5472.14	f_c-4f_m	f_c+4f_m	7472.14
5	7090.17	f_c-5f_m	f_c+5f_m	9090.17
* lower sideband frequencies are the absolute value				

Table 1. Shaded cells show the four low-order partial frequencies that are powers of φ when both the carrier and modulating frequencies are powers of φ (but not equal).

also powers of φ, see Table 1. This unique attribute caught my attention, as this is not the case with √2 or any other irrational number that I am aware of.

3.2 The Golden ratio and the Pitch Space

I then "discovered" [4] that powers of φ were related in the same way as Fibonacci numbers, as seen in Equation 3.

$$\begin{aligned} {}^{n+1} &= {}^{n} + {}^{n\ 1} \\ n &= 1, 2, 3... \end{aligned} \tag{3}$$

Expanding out powers of φ in log frequency results in an equal intervallic division of pitch, as is the case with powers of 2. I have referred to the interval based on this division as a pseudo-octave [7], with an equal tempered division of the pseudo-octave into nine scale steps. I call this the "Stria scale" (StrScl), for the composition in which it was first used.

In three of my compositions I exploited this division of the pitch space and the complementary inharmonic spectra based on the φ and FM synthesis (φFM) as shown in Table 1,

- Stria (1977) — φFM spectra, [9]
- *Phoné* (1981)—harmonic spectra of synthesized singing voice mixed with φFM spectra,
- Voices (2005, v.3 2011)—harmonic spectra of soprano's voice mixed with φFM spectra and synthesized singing voice.

Together with a longstanding interest in aspects of Greek mythology and history, especially the Pythia and her origins, the Golden ratio and the Oracle of Delphi came together in Voices for soprano and interactive computer. But on the way, I became fascinated with the singing voice.

3.3 The Singing Voice: *Phoné* and Voices

In 1978 Jean-Claude Risset invited me to spend a year at IRCAM. Based on Michael McNabb's demonstration in Dreamsong (1978) that capturing the fundamental frequency (phonation frequency) of a sung female vowel tone through time, is to capture the signature of the singing voice, even if it is a sine wave, I set about to synthesize the singing voice with FM synthesis. Taking advantage of McNabb's important insight and Johan Sundberg's vast knowledge of the science of the singing voice, I profited greatly from his presence at IRCAM and was able to synthesize a number of sung vowel tones.

By setting the modulation frequency at the phonation frequency (pitch frequency) and the carrier frequencies at the closest harmonics to a given vowel's formant frequencies, I successfully modeled the target spectrum, as shown in Figure 3. The relationship of the spectral model to the signal generation can be seen in Equation 4. With an appropriate mix of a piece-wise linear random func-

[4] This was a "discovery" in that in 1974, I knew that the ratio between consecutive numbers of the Fibonacci sequence were an approximation of φ, but I had no knowledge of the same relationship between the powers of φ.

tion and a periodic sinusoidal function to approximate the micro-modulation of pitch (phonation frequency) through time, the simulations were convincing. This work is described in "Synthesis of the Singing Voice by Means of Frequency Modulation" [10].

$$e = A_1 \sin(2\pi f_{c1} t + I_1 \sin 2\pi f_m t) + A_2 \sin(2\pi f_{c2} t + I_2 \sin 2\pi f_m t) \quad (4)$$

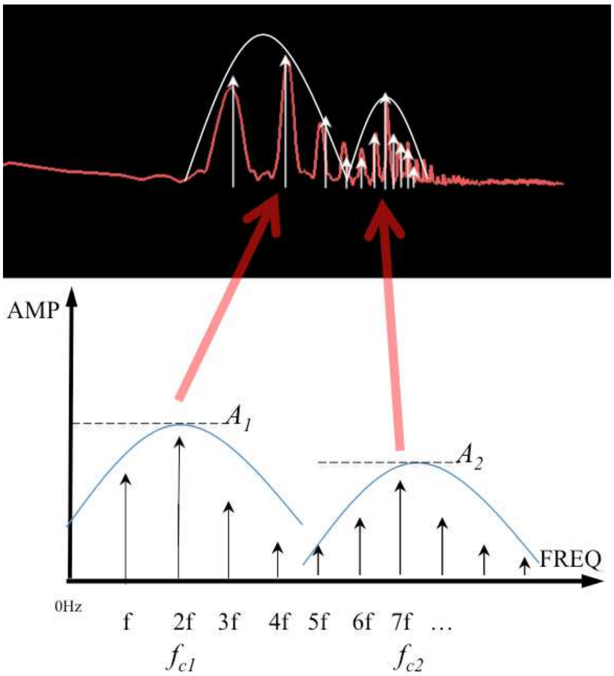

Figure 3. Spectral modeling of the singing voice (or any sound having prominent resonances) can be realized by setting the carrier frequencies, f_{c1} and f_{c2} at the harmonic frequencies, 2f and 7f, closest to the resonance peaks. The target spectrum in red, was captured by sndpeek. Bandwidths of the resonances (blue curved lines) are determined by the indices I_1 and I_2, here ≈ 1.0.

One might ask why synthesize a singing voice when one can sample and then process a real voice? One answer lies in the kind of control one has over the details of the sound material. With synthesis, sound can be formed in ways that are not possible in transformations of sampled sounds.

John Pierce's Eight-tone Canon (1966) [11] could only have been realized by synthesis because the timbres are composed of precisely arranged partials that are ordered but not in the harmonic series. So, too, in Jean-Claude Risset's Mutations (1969), where a set of pitches is heard first as melody, then as harmony and finally folded into timbre [7]. It is the last stage which, again, is composed of precisely tuned partials from the set of pitches that gives an inharmonic, gong-like sound an ineffable quality of sounding "imprinted" pitches.

It was Mutations that inspired me to extend Risset's powerful idea to another level of control based on my research with the singing voice and perceptual fusion [12]. *Phoné* was premiered at IRCAM in 1981.

Over several years I developed the SAIL[5] code around the idea of continuous transformations of sounds through detailed control of the partials and the conditions in which they cohere, or fuse, to be perceived as a single source rather than individual partials. As noted above, Risset demonstrated in Mutations that sinusoids that begin together with amplitude envelopes that are exponential in shape and fall off in duration with increased pitch height, sound gong- or bell-like, but imbued with harmony. The onset of such a tone is shown in Figure 4.

Extending this process to another level of complexity in *Phoné*, each of these sinusoids is the f_{c1} of a two carrier FM process as shown in Equation 4. The amplitude envelopes A_1 do not decay to 0, but rise and are joined by the other three components of the Equation 4, A_2, I_1 and I_2, as the micro-modulation is faded into the mix— a smooth transformation to multiple singing voices.

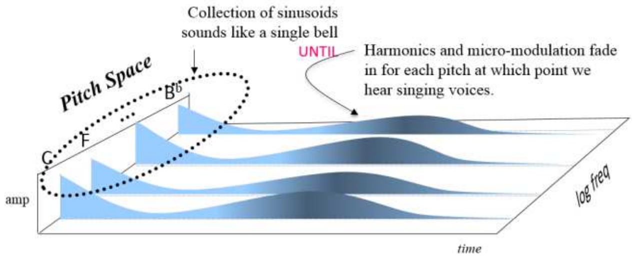

Figure 4. A collection sinusoids with frequencies from the pitch space sound like a bell at the onset. Continuing, they each become a harmonic in singing voice tones, where the change in hue represents the additional harmonics.

Voices makes use of synthesized sounds only and the amplified and processed sound of a soprano. The sounds and pitches are based upon φFM spectra and the StrScl (and its pseudo-octave). The question at the outset was whether or not a well-trained singer could comfortably sing in an unfamiliar spectral complex and in an artificial tuning system? (Details of how the piece was composed have been previously described [7].) The answer seems to be yes and I have found independent confirming evidence as to why this may be so.

4. PARTIALS AND TUNING

Hiding (from me, at least) in the ever increasing corpus in the hearing sciences, is a demonstration CD that has an astonishing (to me, at least) and relevant example that shows the importance of the complementary relationship between spectral space and pitch space. It is astonishing partly because the example is not cast in the context of new music, where it is often difficult to make critical, objective judgments because both material and context are unfamiliar. This example is a synthesized Bach chorale [13], without artifice, where the tones are composed of partials produced by individual oscillators, the ampli-

[5] Stanford Artificial Intelligence Language is a procedural language developed at Stanford in the 1960s-70s. The *Phoné* code was derived from the Stria code in the same language.

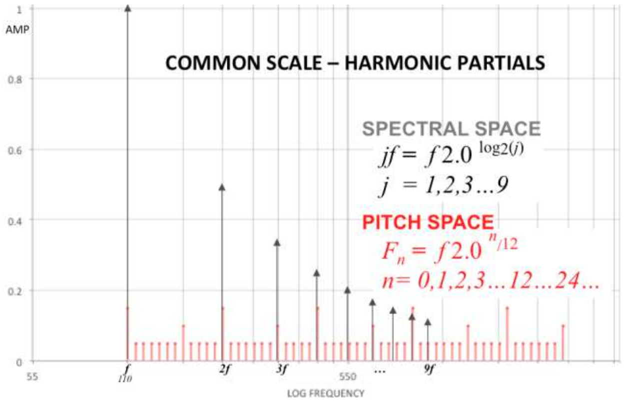

Figure 5.1

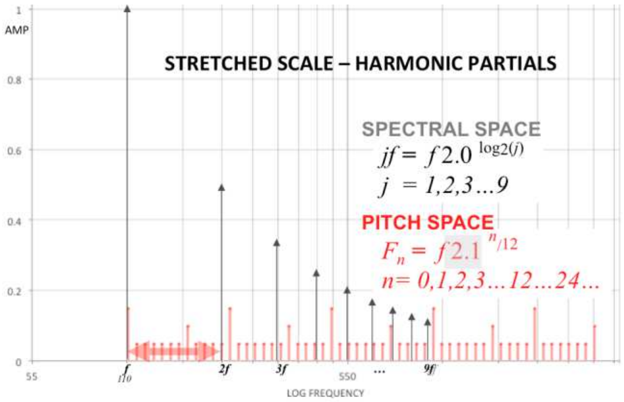

Figure 5.2

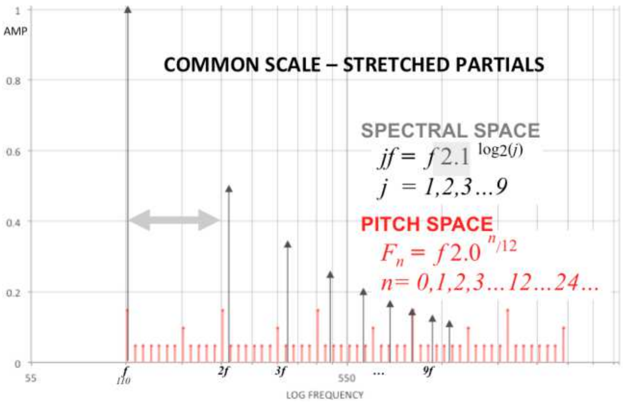

Figure 5.3

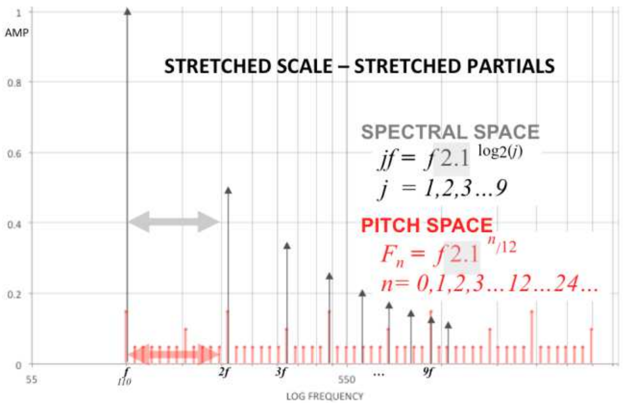

Figure 5.4

tudes of which are similar to those of a sawtooth wave. However, it is not a sawtooth wave and could not be!

The chorale is presented four times where each iteration sounds a different relationship of tones and tuning. The spectral/tuning renderings of the chorale are represented in Figure5.1-4 by a tone having a pitch frequency of 110Hz, where the red colored equation and division along the x axis stand for the pitch space scale and the gray equation and grey partial components their frequency relation to the pitch space.

- In Figure 5.1 the base of both equations is 2.0.
- In Figure 5.2 the base of the pitch equation is increased by 10% to 2.1.
- In Figure 5.3 the base of the spectral equation is increased by 10% to 2.1.
- In Figure 5.4 the base of both in increased by 10%

The 1st corresponding sound example sounds as expected, simple and synth-boring. The 2nd and 3rd sound examples sound out-of-tune, again, as expected. But the 4th example, where both tuning and partials are stretched was not as expected. I had expected it to sound out-of tune, but in a different way than the previous two. In, fact it sounded good, surprisingly— more interesting that the 1st sound example!

When I formed the theoretical underpinnings for Stria and began the time-consuming sound realization, I had wondered if its lissome sound surface was unique because of its φFM spectra? And so with *Phoné*. Engaging a soprano in Voices was a special challenge, because I was unsure how the digital precision of synthesis would interact with the suppleness of a real singing voice. But again, the piece is built on the same "plinth" as Stria and *Phoné*. Finding the Tones and Tuning with Stretched Partials [13] example pointed toward, and gave weight to, a generalization: building sound structures where pitch space and spectral space are complementary may open to an entirely new soundscape.

5. CONCLUSIONS

Understanding the implications of Mathews' diagram freed musical ideas that led me into a field of study, research and creation that I could not have anticipated. The Golden ratio fell into my "ear lap" simply because it was "in the air"— in the culture of the 1970s with M.C. Escher t-shirts, computer graphics and D. Hofstadter's *Gödel,* Escher, Bach: An Eternal Golden Braid.

Much of my inspiration is close to the bits and bytes of sound, the spectral-temporal detail—and to the programming language itself, abstract and cool in its generality, but often provocative and animating when engaged.

6. REFERENCES

[1] C. Shannon, "A mathematical theory of communication." ACM SIGMOBILE Mobile Computing and Communications Review 5, no. 1, pp. 3-55, 2001.

[2] H. D. Lüke, "The origins of the sampling theorem," *IEEE Communications Magazine* 37, no. 4, pp. 106-108, 1999.

[3] M.V. Mathews, "The Digital Computer as a Musical Instrument," Science, Vol. 142, No. 3592, pp. 553-557, 1963.

[4] J. Tenney, "Sound-generation by means of a digital computer," *Journal of Music Theory,* pp. 24-70, 1963.

[5] J-C Risset, and M.V Mathews, "Analysis of Musical-Instrument Tones," Physics Today, vol. 22, no. 2, 23-30, 1969.

[6] J. Chowning, "Turenas: the realization of a dream." *Proc. of the 17es Journées d'Informatique Musicale, Saint-Etienne, France,* 2011.

[7] J. Chowning, "Fifty Years of Computer Music: Ideas of the Past Speak to the Future." *Computer Music Modeling and Retrieval. Sense of Sounds*. Springer Berlin Heidelberg, pp. 1-10, 2008.

[8] Euclid, and D. Joyce. *Euclid's elements*. Clark University, Department of Mathematics and Computer Science, 1998.

[9] O. Baudouin, D. Dahan, M. Meneghini and L. Zattra, The Reconstruction of Stria, Computer Music Journal, Vol. 31, Num. 3. MIT Press, Cambridge, 2007.

[10] J. Chowning, "Synthesis of the Singing Voice by Means of Frequency Modulation," in Sound Generation in Winds, Strings, Computers. Royal Swedish Academy of Music, 1980, No. 29. Reprinted in Current Directions in Computer Music Research, Edited by M. Mathews and J. Pierce, MIT Press, 1989.

[11] J. Goebel, (producer). The Historical CD of Digital Sound Synthesis. Computer Music Currents 13, Schott Wergo, 1995.

[12] J. Chowning, "Perceptual Fusion and Auditory Perspective," in P. Cook (ed.), Music, Cognition, and Computerized Sound: An Introduction to Psychoacoustics, Cambridge, MA: MIT Press, 1999, pp. 261-275.

[13] J. Bach, "Als der gütige Gott," in Tones and Tuning with Stretched Partials, Auditory CD, Acoustical Society of America, 1987, No 31. (http://asa.aip.org/discs.html)