Fifty Years of Computer Music: Ideas of the Past Speak to the Future

John Chowning[1]

[1] CCRMA, Department of Music, Stanford University, Stanford, California 94305
jc@ccrma.stanford.edu

Abstract. The use of the computer to analyze and synthesize sound in two early forms, additive and FM synthesis, led to new thoughts about synthesizing sound spectra, tuning and pitch. Detached from their traditional association with the timbre of acoustic instruments, spectra become structured and associated with pitch in ways that are unique to the medium of computer music.

1 Introduction

In 1957, just fifty years ago, Max Mathews introduced a wholly new means of making music. An engineer/scientist at Bell Telephone Laboratories (BTL), Max (with the support of John Pierce, who was director of research) created out of numbers and code the first music to be produced by a digital computer. It is usually the case that a fascination with some aspect of a discipline outside of one's own will quickly conclude with an experiment without elaboration. But in Max's case, it was the beginning of a profoundly deep and consequential adventure, one which he modestly invited us all to join through his elegantly conceived programs, engendering tendrils that found their way into far-flung disciplines that today, 50 years later, continue to grow without end.

From the very beginning Max's use of the computer for making music was expansive. Synthesis, signal processing, analysis, algorithmic composition, psychoacoustics—all were within his scope and all were expressed and described in great detail in his famous article [1] and the succession of programs MUSIC I-V.[1]

It is in the nature of the computer medium that detail be elevated at times to the forefront of our thinking, for unlike preceding music technologies, both acoustic and analogue, computers require us to manage detail to accomplish even the most basic steps. It is in the detail that we find control of the sonic, theoretical and creative forms. And it is through paying attention to detail that we reveal our scientific/engineering insights or our artistic expression—our own voice.

The first examples of computer-generated music produced by Max and by John Pierce at BTL were rich in ideas, including algorithmic composition, novel tuning, matching tuning systems to complementary spectra, imaginative and compelling

---

[1] For a complete account of Max Mathews' work and publications, see http://www.ina.fr/produits/publications/collections/collec_11.fr.html.

graphics and visualizations and, soon following, controllers [2]. It is fortunate that these two scientists/engineers—who cultivated a nexus between science and art, and who invited many composers and artists to their laboratories (e.g., Varèse and Cage) to share the possibilities that they saw—were willing to place these nascent musical studies in the public view, confident in the intellectual content of their ideas, which few others could see. Some of their ideas remain as compelling now as they were then and should be "re-viewed" given the enriched domains of application at this 50-year mark.

2 Breakthroughs

The richness of the ideas in these early examples was not matched by the quality of the sounds with which they were expressed. Little was known about some important aspects of perception and the acoustics of musical instruments. Two important composers were invited by Max to work at BTL, both of whom made important contributions in this area in addition to creating compositions: James Tenney and Jean-Claude Risset.

Preceding Max's famous article by a few months was an article by Tenney that described in exquisite detail the program that Max had created [3]. Tenney had been invited by John Pierce and Max to work at BTL beginning in 1961. He had studied with the visionary Lejaren Hiller at the University of Illinois, so he came prepared in matters of programming and stochastic processes in composition. During his three years at BTL he made several important contributions; he created compositions using this new medium, and he wrote in great detail about what he had learned from Max and how Tenney had constructed his compositions. Because he was a composer, Tenney's description of Max's MUSIC IV was from a musical view, and it remains an exemplar of clarity and completeness.[2]

But, important to the points being presented in this paper, he came upon a music-driven question in his compositions using MUSIC IV for which there was no answer, so with Max's guidance, he did a study regarding the *perception* of attack times [3]. The italicization is to draw attention to two points: 1, the fact that from the outset psychoacoustics had been seen by Max as one of the crucial disciplines in the advancement of computer music[3] and 2, that musicians have a particular sensitivity to details of auditory perception.

---

[2] His early interest and important contributions notwithstanding, Jim Tenney did not continue in computer music, but rather became a distinguished teacher, performer and composer of acoustic music. He died August 24, 2006.

[3] Max wrote in 1963 "At present, the range of computer music is limited principally by cost and by our knowledge of psychoacoustics [4]."

2.1 Risset Uncovers the Microstructure

There is no doubt that the most important breakthrough in the early days of computer music occurred when Jean-Claude Risset and Max began detailed computer studies in the analysis, synthesis and perception of acoustic instrument tones, culminating in Risset's *An Introductory Catalogue of Computer Synthesized Sounds* [1]. With this work the medium of computer music reached a level beyond Max's correct but abstract assertion that computers (coupled with loudspeakers) can produce any perceivable sound. The capability of simulating natural-sounding tones presupposes an understanding of the perceptual relevance of the physical stimuli, only some of which have been "selected" as meaningful by the auditory system.

2.2 FM Synthesis—40 Years

It was forty years ago that this author 'stumbled" upon FM synthesis [4]. The actual date is not known. Not having a scientific or engineering background, I did not have the habit of keeping dated lab notes, but I did keep notes. There is a record of my having visited BTL on December 18, 1967 when I showed the data that I used in my first trials to Max, Risset and Pierre Ruiz and played for them the examples. It was a month or two before, almost certainly late at night, while experimenting with extreme vibrato frequencies and depths that I realized "there is more here than at first meets the ear."

Its discovery was not a purposeful search—that is, stemming from a realization, from looking at the equation, that there might be some interesting experiments to try—rather, it was altogether a discovery of the "ear."

One must remember that while the theoretical potential for the production of rich dynamic sounds with the computer was great, the knowledge required for realizing this potential was meager. Risset's catalogue was in progress and little known outside of BTL. Furthermore, the cost in computer time was enormous, limiting the complexity of synthesis algorithms. Deep into the details of digital reverberation at the time, I was keenly aware of this issue. My "ear" was continually scanning for any sound having internal dynamism, coupled oscillators, random vibrato, etc. That I found it within such a computationally efficient algorithm was certainly partly chance, but then I was also certainly prepared for that chance.

The first experiments were each only a few seconds duration, because of the tens of minutes of compute time on a time-shared system. But they do show that from the outset, all of the essential features were noted that would eventually be developed and used in musical contexts:

- both harmonic and inharmonic spectra could be produced
- a change in frequency deviation (Δf) produced a change in bandwidth of the spectrum
- the spectrum is conserved through the pitch space with a constant ratio of FM frequencies

As it turned out, these parameters of FM synthesis have a remarkable perceptual relevance.

As mentioned above, Risset's study of trumpet tones had a major influence on my own development of FM synthesis. I first heard about this study on the aforementioned visit to BTL in 1967, during which I showed my first experiments in FM synthesis. Risset explained his analysis and re-synthesis of trumpet tones and played some examples. It was not until 1970, however, that I fully appreciated the importance of his discoveries about trumpet tones.

While working on the FM synthesis of percussive sounds, I noted that in nearly all tones of this class the amplitude envelope and the envelope controlling the modulation index were very similar if not identical. I also noted that there was as strong a correlation of the perception of ‘strike force’ to the modulation index as there was to intensity. I considered other classes of tones where this might be the case, and I remembered Risset’s explanation of the “signature” of trumpet tones, some three years previous. With only a few attempts I was able to create credible brass-like tones by simply coupling a single function to the amplitude and modulation index envelopes with appropriate scaling. I realized that this correlation of force or effort (strike force, breath and bow pressure velocity, etc.) to the bandwidth and/or high-frequency emphasis of partials can be generalized to all natural sound and that the parameters of FM synthesis provided a straightforward implementation of this important correlation[4].

Then began a rapid development of FM synthesis[5], and the eventual licensing of the technology by Stanford University to Yamaha—the rest is history.

3 Structured Spectra and Pitch Space

There are two ways in which additive synthesis and FM synthesis have been used that merit emphasis, because they touch upon issues that are important beyond any particular means of synthesis. John Pierce and Max foresaw one way in the early years: the creation of a non-traditional scale that has a structural link to timbre, where the frequency ratios from the scale are used in the construction of the tone’s spectra. Karlheinz Stockhausen created a similar relationship between pitch and spectrum in his Studie 1 (1953). Risset, however, used synthesis in a manner not foreseen—a manner imaginative and evocative.

3.1 Constructing Spectra ***in*** the Pitch Space

The final example in Risset’s catalogue stands as a striking advance in computer music, although little recognized and little exploited. It is the first instance where pitch is used to express timbre in the same functional manner that pitch expresses melody and harmony, that is, melody-harmony-timbre all within the pitch space.

---

[4] The ease with which spectral change could be coupled to effort (key velocity) is one of the reasons for the YAMAHA DX7's remarkable success.

[5] The first real-time FM synthesis was programmed on a DEC PDP-15 computer by Barry Truax in 1973, while studying in Utrecht. At Stanford, Bill Schottstaedt developed a particularly powerful form of the algorithm that was used in many compositions for many years.

Pitch is composed sequentially as line and simultaneously as harmony, for which there are rich functional theories, but composing timbre as a collection of partials drawn from the pitch space cannot be achieved with acoustic instruments and falls squarely in the domain of computer music.

The sound potential of any instrument is vast, but limited—the partials that make up an instrument's tone can only be partly modified by performance techniques and devices such as mutes. A clarinet and a violin can play the same pitch at the same loudness for the same duration, but they cannot be made to have the same spectrum through time—the frequency and intensity of an instrument's partials are locked within boundaries defined by its and the performer's physical properties.

Risset realized in his timbre studies that in creating natural sounding complex timbres by summing numbers of sinusoids (pure tones) where each sinusoid can have its own independent control over intensity and frequency through time, he had unlocked timbre from any physical constraints. He could create tones that cannot exist in the natural world, complex timbres where the partials themselves are a part of the pitch space. He composed a short pitch sequence that is heard first sequentially in time (melody), then simultaneously in time (harmony), and then again simultaneously with exactly the same pitches but now as partials associated with a single sound source, as shown in Fig. 1. [2].

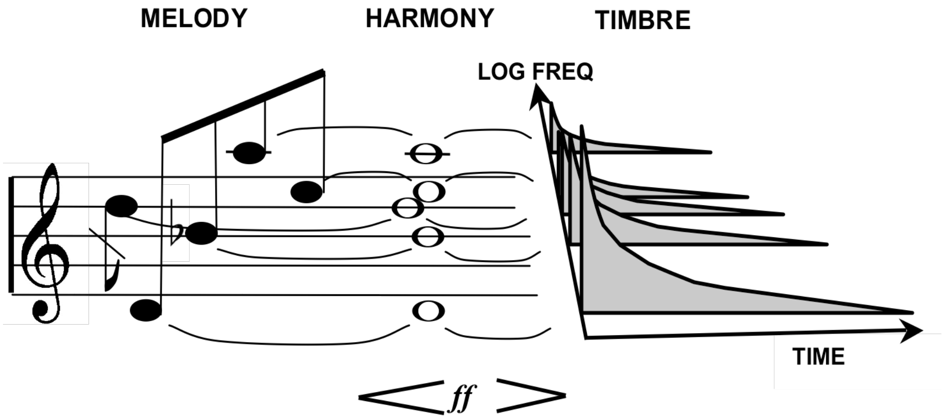

Fig. 1. Pitches become the partials of the gong-like tone, linking timbre to the pitch space in a manner uniquely possible with computers- from *Mutations* 1969 Jean-Claude Risset

Because all of the partials die away in a similar manner, they fuse and are heard as timbre rather than harmony. The timbre is similar to that of a gong, but a gong whose spectrum is imprinted with pitch information, giving the sound an extranatural structural link to the preceding. Risset's was an altogether new conception, uniquely possible with computers, and beautifully framed in several of his compositions, first in *Mutations* (1969).

3.2 Constructing Spectra *and* the Pitch Space

John Pierce and Max saw early on that using the computer for both control and synthesis could unlock tuning systems from physical constraints, just as Risset had unlocked timbre. Max composed a piece, *The Second Law* that is entirely made up of noise, entirely free of common understandings of pitch, yet expressing pitch. In his *Eight-Tone Canon* (1966)[2], Pierce divided the octave into eight equal steps—the even-numbered steps (equal to the multiples of three in a twelve-step division) and odd-numbered steps each form a diminished seventh chord. But what is interesting about this short piece is that Pierce used tones composed of sums of sinusoids that progress from octave to half-octave to quarter octave, with each iteration of the canon. Except for the octave, the spectra are inharmonic, but composed of frequencies that are common to the pitch space!

***Stria* (1977.** While the above example is not rich in the sonic sense, it is a compelling and powerful idea that I found especially evocative because of my interest in spectra composed of *ordered* inharmonic partials —some of which are simply produced by FM synthesis. One class of such spectra that I found particularly interesting is based upon carrier-to-modulator frequency ratios (f_c/f_m) derived from the Golden Ratio or $\Phi \approx 1.618$. Remembering Pierce's canon, I conceived a composition in the mid 1970s that is based upon spectra structured in a way that is complementary to the division of the pitch space. The traditional octave is replaced by a pseudo-octave based upon powers of the Golden Ratio (Φ^n) rather than powers of 2, and the spectra are produced by values of f_c/f_m that are also powers of Φ as can be seen in Fig. 2.

After several years thinking about its theoretical underpinnings, I realized *Stria* in the months from July to October 1977. The composition of the work was dependent upon computer program procedures, specially written to produce the enormous amount of data that specified the details of the complementary relationship between pitch space and the ordered inharmonic partials. In addition, these procedures are at times recursive allowing musical structures that they describe to include themselves in miniature form - similar in idea to the embedded fractal geometries of Mandelbrot. From the beginning, *Stria* softly unfolds element by element, overlapping such that the inharmonic partials create increasing spectral density, ordered by ratios of Φ in both time and pitch. The major division of STRIA is at the Golden Section where recursion is used to create enormous acoustic mass. The final section of the composition is the inverse of the beginning, becoming ever less complex until it ends with a fading pure tone

Stria was first presented on October 13, 1977 at the Centre Pompidou as part of IRCAM's concert series "La Voix des voies" produced by Luciano Berio. The composition is fully described in the Fall and Winter issues of the Computer Music Journal, 2007 [6].

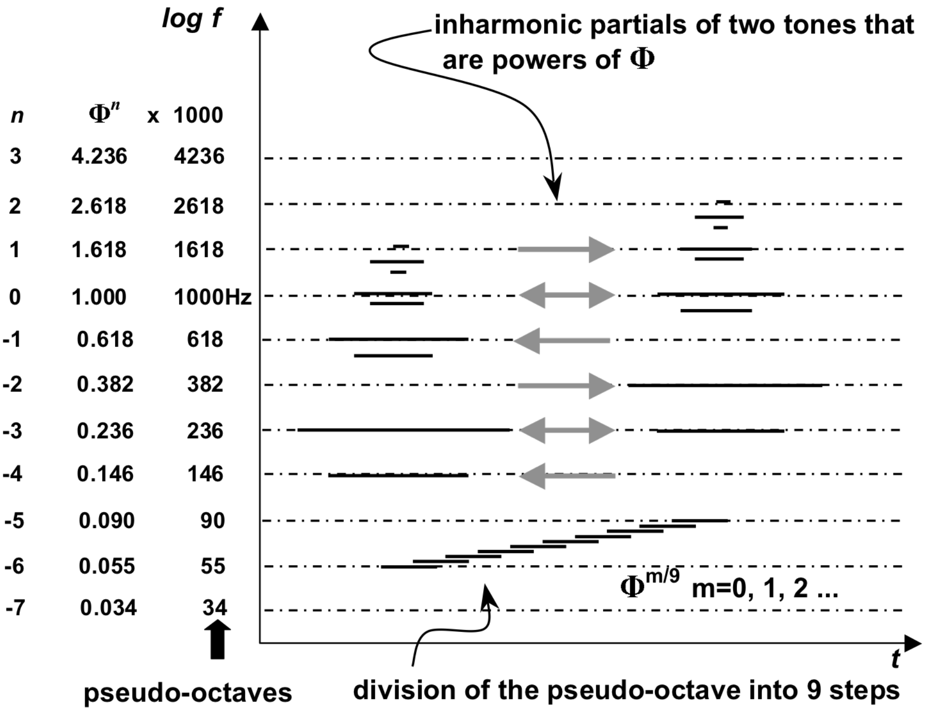

Fig. 2. The pitch space in both *Stria* and *Voices* is based upon pseudo-octaves that are powers of the Golden Ratio Φ, which are further divided into 9 steps. The spectra of tones generated by FM c:m ratios that are also powers of Φ, contain partials that are powers of Φ, here shown at the interval of a pseudo-octave

***Voices* v.2 2007.** *Voices,* for soprano and interactive computer, uses the same division of the pitch space and structured spectra as in *Stria*. Again, all the sounds are produced by FM synthesis and all the spectra are generated from ratios based on Φ, as noted above (except for a few instances of voice-like tones that use integer ratios). The formal structure of *Voices* is altogether different, however, and requires a larger set of differentiated sounds than did *Stria*.

The important and initial question was how well would a soprano, both as a performer and as a "sound," fit into this 'artificial' pitch/spectral space, where first, the scale is unfamiliar to the performer and not related to any of the common modes or tunings and second, the partials of sung vowel tones are harmonic and do not share the same spectral distribution? Can one mix such a sonic artifact, totally dependent upon the computer for its existence, with a natural, perhaps the most natural, musical sound, the singing voice?

The music performance problem would seem to be a major hurdle for the soprano. While a single scale step in the 9 step/pseudo-octave division only differs from the traditional semi-tone by 7 cents, the maximum difference in the progression is nearly a quartertone and there are 13 steps in the interval closest to a true octave, as shown in Table 1.

Table 1. A comparison of the scale used in *Voices,* $\Phi^{n/9}$, with the common tempered scale, $2^{n/12}$, showing the closest scale degree and the difference in cents. The 6th step of the common scale repeats to maintain the proximate *Voices* step and show that it is the 13th step of the *Voices* scale that is closet to a true octave

n	Voices Scale	Pitch	Cents	Common Scale	n
0	1.000	a		1.000	0
1	1.055	a#(-)	-7	1.059	1
2	1.113	b(-)	-14	1.122	2
3	1.174	c(-)	-22	1.189	3
4	1.238	c#(-)	-29	1.260	4
5	1.306	d(-)	-37	1.335	5
6	1.378	d#(-)	-44	1.414	6
7	1.454	d#(+)	48	1.414	6
8	1.534	e(+)	41	1.498	7
9	1.618	f(+)	34	1.587	8
10	1.707	f#(+)	26	1.682	9
11	1.801	g(+)	19	1.782	10
12	1.900	g#(+)	11	1.888	11
13	2.004	a(+)	4	2.000	12

It is my good fortune to have had a soprano[6] at hand with whom I could work during the initial stages of the composition to test my hypothesis: singing in this unusual scale is possible if the structured inharmonic spectra of the accompanying tones are infused with complementary pitch information, since most good performers tune to context. I included in the program (written in MaxMSP) the option for the singer to give herself a cue tone for the current target pitch or the following target pitch. In fact, the option is rarely used since the singer seems to easily tune to the partials of the structured spectra—to the context, as hypothesized.

The other part of the initial question, how well would the soprano sound, having partials in the harmonic series, fit within a context composed of dense inharmonic partials, albeit structured? The somewhat surprising answer: the performer and the listener are unaware of any spectral mismatch. Moreover, one senses an overall pitch coherence that is more like a soprano singing with an acoustic instrument ensemble having harmonic spectra, than singing with idiophones having dense inharmonic spectra such as gongs and bells. There are several possible reasons that there is no overall percept of “out-of-tuneness” or psychoacoustic dissonance.

---

[6] Maureen Chowning, for whom *Voices* is written, has had experience singing in alternative tunings, e.g. *Solemn Songs for Evening* by Richard Boulanger, written in the Pierce-Bohlen scale.

While the spectra of low tones in *Voices* are often dense, they are selectively dense with partials of a single tone spaced at intervals from large to small (in log frequency), as is the case with the harmonic series. In addition the spectra are composed such that the energy is concentrated around the low order partials (small modulation index). Therefore, whether or not low order partials fall within a critical band is dependent upon the interval of two tones within the pitch space, as is the case with partials in the harmonic series.

Another reason is that the concentration of harmonic energy in typical soprano tones is limited to the low order harmonics, especially the fundamental, which reduces the incidence of strong partials interacting within critical bands.

Finally, there is a third reason, which is somewhat speculative, why there is little psychoacoustic dissonance. Critical band theory is based upon perceptual experiments using stimuli having few variables and partials that are highly stable, quite unlike sounds of the natural listening experience. We know that the auditory system responds to partials in a different way when mediating temporal factors are present such as amplitude envelopes or synchronous micro-modulation (e.g. random or quasi-periodic vibrato).

This internal dynamism brings into play an additional perceptual theory based upon grouping and common fate from the Gestalt laws of perceptual organization. It is a higher-level mechanism (probably not in the peripheral auditory system) that causes partials to fuse or cohere, where individual partials are difficult or impossible to distinguish, and they become identifiable as a source, known or unknown, and segregable within a collection of sources.

It seems that these temporal features, which are intrinsic to our "out of lab" perceptual experience, may reduce the importance of the interaction of individual partials relative to critical bands because the interaction becomes transitory and no longer stable.

Dynamic partials not only animate the sounds of which they are a part, harmonic or inharmonic, but contribute to the surface allure of the larger sound context, leading the ear through time in a complex of detailed multi-dimensional spaces of timbre, location, loudness and—most importantly regarding dissonance—pitch. The pitch space is loosened from its abstract skeletal form by the internal dynamic detail, and accommodates the sound of the soprano, whose fluid expression derives from its own internal dynamism.

4 Conclusion

During the thirty years since composing Stria, I have often wondered whether the integrated spectral tuning and pitch tuning worked because of the particular attributes of the work itself, the manner in which the work slowly unfolds from sparse to dense spectra and having no other spectral forms than those rooted in the Golden Ratio? Is it a pitch/spectrum construct that is unique to the piece?

My experience with *Voices* suggests that it is not. It could it be that these early ideas—Risset's structured spectra, Max's and Pierce's joining of odd tunings with complementary spectra and Max's evolving pitch space in his melodic metamorphoses, can be generally exploited with synthesized sound, with physical-models where the "physical" is infinitely malleable, or even sampled sounds, especially with the availability of new stable high-Q filters [7]. A medium is defined by its distinctive attributes and these ideas are certainly unique to music made with computers.

References

1. Mathews, M. V. :"The Digital Computer as a Musical Instrument," *Science*, Vol. 142 (1963) No. 3592, 553-557
2. Goebel, J. (producer): The Historical CD of Digital Sound Synthesis. Computer Music Currents 13, Schott Wergo (1995)
3. Tenney, J.: Sound-Generation by Means of a Digital Computer, Journal of Music Theory, Vol.7, No.1, Spring 1963) (Available http://links.jstor.org/)
4. Chowning, J.: The Synthesis of Complex Audio Spectra by Means of Frequency Modulation, *J. Audio Eng. Soc*. 21, 7 (1973)
5. Chowning, J.: Composing the Sound Itself, *John Chowning Portraits Polychrome*,. Nouvelle edition augmentée. Institute national de l'audiovisuel, Paris (2007)
6. Meneghini, M.: An Analysis of the Compositional Techniques in John Chowning's *Stria,*
   Zattra, L.: The assembling of *Stria* by John Chowning: A philological investigation,
   Baudouin, O.: A Reconstruction of *Stria,*
   Dahan, K.: Surface Tensions: Dynamics of *Stria*
   The Reconstruction of Stria. Computer Music Journal, Vol. 31, Num. 3. MIT Press, Cambridge (2007)
7. Mathews, M.V., Smith, J. O.: Methods for synthesizing very high Q parametrically well behaved two pole filters,' Proceedings of the Stockholm Musical Acoustics Conference (SMAC-03),http://www.speech.kth.se/smac03/, (Stockholm), Royal Swedish Academy of Music, Aug. 2003,available online, with sound examples, at http://ccrma.stanford.edu/~jos/smac03maxjos