MANIFOLD COMPOSITIONS, MUSIC VISUALIZATION, AND SCIENTIFIC SONIFICATION IN AN IMMERSIVE VIRTUAL-REALITY ENVIRONMENT

Hans G. Kaper
Senior Mathematician
Mathematics and Computer Science Division
Argonne National Laboratory
Argonne, Illinois 60439-4844, USA.
email: kaper@mcs.anl.gov
http://www.mcs.anl.gov/people/kaper/index.html

Sever Tipei
Professor, Manager Computer Music Project
University of Illinois at Urbana-Champaign
1114 W. Nevada St.
Urbana, Illinois, 61801, USA.
email: s-tipei@uiuc.edu
http://cmp-rs.music.uiuc.edu/people/tipei/index.html

Abstract:
An interdisciplinary project encompassing sound synthesis, music composition, sonification, and visualization of music is facilitated by the high-performance computing capabilities and the virtual-reality environments available at Argonne National Laboratory. The paper describes the main features of the project's centerpiece, DIASS (Digital Instrument for Additive Sound Synthesis); "A.N.L.-folds", an equivalence class of compositions produced with DIASS; and aplication of DIASS in two experiments in the sonification of complex scientific data. Some of the larger issues connected with this project, such as the changing ways in which both scientists and composers perform their tasks, are briefly discussed.

1. Introduction
Music composition, sound synthesis, sound visualization, and scientific sonification are the four elements brought together in an ongoing interdisciplinary research project at the University of Illinois at Urbana/Champaign (UIUC) and Argonne National Laboratory (ANL). The glue binding these diverse elements is DIASS (Digital Instrument for Additive Sound Synthesis), a software system developed jointly at the two institutions and implemented on the high-performance computer architectures of Argonne's Center for Computational Science and Technology (CCST).

The expertise required in the project ranges from music theory to scientific computing and computer science. The principal investigators are an applied mathematician and a composer, both with extensive experience in sound space, who have access to experts in computer science.

1.1. Manifold Compositions
The music composition part of this project is based on the idea that the computer is a composer's collaborator. While the human artist controls the overall outlook of the music and its abstract structure, the machine provides the details of the piece and ensures a certain degree of randomness unencumbered by cultural conditionings [Tipei, 1987]. By supplying a set of rules and initial conditions, the composer sets in motion a well-defined process but does not interfere with the process or its end result.

A "manifold composition" consists of all actual and potential variants of a musical work that is composed with the assistance of a computer and uses the same code and the same set of input data, but contains elements of indeterminacy [Tipei, 1989]. Since the details of a piece depend on random occurrances, the number of variants of a manifold composition is unlimited. The variants of a manifold composition share the same formal structure, pitch, and rhythmic material and have similar textures, while variations can range from a slight rearrangement of the notes in the score or the sounds on the tape to a radical alteration of the textures and even of the succession of sections in the piece.

In a way, a manifold composition is similar to the serigraphs in the visual arts, except that individual members of the composition class are usually more distinct from one another. There is also a similarity between manifold compositions and aleatory music, which was introduced four decades ago and described by Umberto Eco as "the actualization of a series of consequences whose premises are firmly rooted in the original data provided by the author" [Eco, 1979]. It is useful to remember that, from the beginning, some composers of aleatory music have referred to it as "fields of possibility" [Pousseur, 1958] and "stochastic composition of fields" [Stockhausen, 1959]. Similarities as well as differences between manifold compositions and aleatory music have been pointed out in detail [Tipei, 1989], along with their connection to the concept of "open work" [DeLio, 1984; Eco, 1979; Interface, 1987].

Manifold compositions are the product of a well defined aesthetic that assumes a speculative and experimental attitude on the part of the composer. According to this aesthetic, music deliberately promotes a world view and is concerned with abstract concepts, not with mundane and anecdotal details; as an artistic tool of investigation, it asks questions about our existence and our place in the world. In doing so, it employs a logic and materials spawned by its own content instead of prefabricated forms and jargons. Its entertainment function is relegated to a secondary role, that of the rhetoric needed to communicate effectively with an audience.

These goals were shared by previous attempts to generate multiple variants of a work by computer (such as the "ST-..." pieces by Xenakis, the "Algorithms" by Lejaren Hiller, G. M. Koenig's "Segmente," and Larry Austin's "Photophorms"). However, most of those attempts involved significant changes in the data between variants, and none of them met the requirement of being mass-produced, which characterizes manifold compositions. The tools available to these composers were simply inadequate to realize large numbers of variants.

1.2. A.N.L.-folds

An example of a manifold composition is "A.N.L.-folds" for computer-generated tape. Seventeen variants of "A.N.L.-folds" were produced as part of the project; five of them appear on a CD with compositions by Sever Tipei, three were performed at the "Funny Music Festival" at the UIUC, July 16, 1997, and two were performed at the "Sound of Israel Festival" of the Hochschule der Kuenste, Berlin, May 20, 1998 as "BERLIN-folds # 1" and "BERLIN-folds # 2". Under the titles "Sonic_0" through "Sonic_50", fifty more variants were performed during "Sonic Residues," a day-long event organized at the Linden Gallery in St. Kilda (Melbourne), Australia, on December 21, 1997.

All "A.N.L.-folds"/"Sonic" variants are exactly 2 minutes and 26 seconds long. They start in the same way - an arpeggio on a low B flat's overtones. Almost half-way through the piece, they all share the same quote, the "Argonne Chime," seven sounds spelling the name of the Laboratory: A, Re, G, sOl, NN (two non-pitched percussive sounds), and E. The ending is a mirror image of the beginning chord. Between these easy to recognize pillars are sections in which all sounds and their attributes (such as start time, duration, pitch, loudness, vibrato, tremolo, reverberation, transients) as well as the overall density of texture are regulated through stochastic distributions. Smooth and seamless transitions between sections are realized by dynamically changing mean values and standard deviations for all distributions to achieve continuity.

There are five such stochastic sections; their durations are always the same, and each is characterized by a particular density, average duration of sounds, and type of effects. Distinct seeds for the random number generator produce variations in the way these and other elements of the music are distributed. A glimpse of this can be caught at the beginning of the piece, where the upper portion of the B flat arpeggio already diverges slightly from variant to variant. The seed also triggers changes in broader and more obvious aspects of the piece. For example, there are three possible choices for a variant's average density of events: low (approximately 2-300 sounds), medium (approximately 3-400 sounds), and high (approximately 4-500 sounds). Each of these choices is also linked to a specific way in which certain sound effects dominate in various sections.

To our knowledge, this type of experiment was never before performed on a scale of this magnitude. Two questions come to mind: How different are these variants? Can the variants be recognized as members of the same equivalence class? After listening to almost six dozens of "Sonic" and "A.N.L.-folds," most of them in one session, we claim that they can without doubt be identified as belonging to the same manifold composition and that they present enough individual features to keep the listener's interest, even when performed successively in large numbers. On a continuous scale having at one end two distinct performances of the same piece of traditional music by the same artist and at the other two works in the same form by different composers, the "A.N.L.-folds" fall somewhere in the middle, but closer to the latter.

2. Software

The "A.N.L.-folds" were meant in part to demonstrate the capabilities of DIASS, our software for digital synthesis. Adopting to a certain extent the MusicN paradigm, DIASS consists essentially of two parts: an editor through which the user enters and modifies the instructions for the instrument (score file) and the instrument proper, which computes the samples. The DIASS instrument functions as part of the M4C synthesis language developed by Beauchamp and his associates at UIUC [Beauchamp, 1993].

2.1. The editor

The editor comes in two flavors, fast and slow, both using essentially the same code. The slow version accepts input through a menu-driven interface (a GUI was written, but had problems with portability) and allows for experimentation and detailed testing, one sound at a time. Once the user has a good idea of the desired types of sound, the fast version, which takes the input data automatically from a script, is employed. The script itself is created by a composition program.

2.2. The instrument

Like all additive-synthesis instruments, DIASS creates complex sounds through a summation of simple sine waves (partials). Unlike other similar instruments, however, DIASS can handle an arbitrary number of complex sounds simultaneously and/or in sequence, each sound can be made up of an arbitrary number of partials, and each partial can be controlled in up to 25 ways [Kriese and Tipei, 1992; Kaper, Ralley, Restrepo, and Tipei, 1995]. Some of the controls are static; they do not change for the duration of the sound. Obvious examples are the starting time, duration, and phase of a partial; less obvious parameters define reverberation features, including the size of the "hall" and the apparent strength of the reverberation. Other control parameters are dynamic; their evolution is governed by specified envelope functions, which may be piecewise linear or exponential. Panning (apparent shifting of sound location), tremolo (amplitude modulation), vibrato (frequency modulation), and the relative importance of transients are examples of dynamically controlled parameters.

The independent controls over each partial ensure an unusual degree of flexibility. On the other hand, partials can also join in to form sounds or complex waves. A set of "macros" performs global operations over given collections of sine waves gelling into a sound. The macros relieve the user from the burden of specifying individual control parameters for each partial. For example, the loudness is specified for an entire sound, and all the component waves are automatically scaled according to a spectrum defined ahead of time. When a glissando is desired, macros ensure that all partials slide at the same pace. Similarly, "tuning" and "detuning" are accomplished by smooth changes of the frequency ratios of a sound's overtones. When applying amplitude or frequency modulation, all waves belonging to the same sound are affected at the same rate; and if a random element is present, all waves access the same random number sequence. The last feature is especially significant because it is an important factor in the perception of a complex wave as one sound, rather than a collection of independent sine waves.

Global operations at the sound level, as encapsulated in the macros, are indeed time saving and, in certain situations, necessary. However, by digitally recreating functions of equipment one found in analog studios - wave generators, filters, mixers, reverberating units, etc. - the macros reinforce the old paradigm on which most traditional sound-synthesis software is based. Behind them is the even more entrenched habit of looking at music as being made up of sounds produced by an orchestra of instruments and read from a score. DIASS indeed accommodates this type of thinking. But in more experimentally inclined minds it can stimulate new ways of looking at sounds, for example by allowing for the possibility of partials changing allegiance and becoming part of a different sound (morphing) or just acting on their own.

2.3. The loudness routines

A unique feature of DIASS is its capability to control the loudness of a sound and synthesize sounds that are perceived by the listener as being "equally loud". The perception of loudness is a subjective experience and depends, among other things, on frequency; this portion of DIASS implements relevant results of psychoacoustic research in software and has been described in more detail elsewhere [Kaper, Ralley, and Tipei, 1996].

The loudness routines enable the user to produce an entire musical work in a single run, even when the sounds cover a wide dynamic range. The "anticlip" macro eliminates the popping noise or clipping that occurs when the computation of a sample results in a number outside the interval (-2**15, +2**15), the largest signed integer that can be represented with 16 bits, by automatically scaling the entire sound. To appreciate the difficulty inherent in this scaling process, consider the case of a sound cluster consisting of numerous complex sounds, all very loud and resulting in clipping, followed by a barely audible sound with only two or three partials. If the cluster's amplitude is brought down to 2**15 and that of the soft tiny sound following is scaled proportionally, the latter disappears under system noise. On the other hand, if only the loud cluster is scaled, the relationship between the two sounds is completely distorted. Many times in the past, individual sounds or groups of sounds were generated separately and then merged with the help of analog equipment or an additional digital mixer. The loudness routines in DIASS deal with this problem by adjusting both loud and soft sounds, so their perceived loudness is equal to the one specified by the user, while preventing overflow.

DIASS was never intended as a real-time instrument. The addition of the loudness routines to the original framework slowed down its performance time considerably. Work is in progress to rewrite the entire program in C++, improve the I/O features, and optimize the code. It is an ambitious overhaul, which has real-time performance as its ultimate goal. Having used DIASS for a number of years, we have a better understanding of our needs and of ways to increase the efficiency of the program.

2.4. Need for high-performance computers

Because of the large number of controls over the behavior of each wave, DIASS is computationally intensive and needs significant amounts of memory. DIASS was conceived as a combination of finesse and power, where the complexity of the task is never sacrificed for the sake of expediency. The idea behind its design was to always use the maximum computing power available, under the assumption that in a not too distant future the same power would be accessible from the user's desktop.

The instrument proper, the engine that computes the samples, comes in a sequential and a parallel version. The two versions use the same code, but the parallel version uses in addition MPI, a message-passing library [Gropp, Lusk, and Skjellum, 1994]. In the case of the "A.N.L.-folds," the sound samples were computed on a 144-node IBM SP. In Argonne's CCST, the SP is a component of the "Quad Machine," which also includes storage, visualization, and media components. Each computational node has substantial RAM and local disk attached to support the operating system and for temporary storage. This summer, the IBM SP was joined by an SGI Origin2000 with 128 nodes, 12 of them Infinity Reality engines.

For the DIASS instrument, parallelism is implemented at the sound level, not at the wave (partial) or sample level, to minimize communication among processors and enable all partials of a complex wave to access the same sequence of random numbers, as required by some macros [Kaper, Restrepo, and Tipei, 1995].

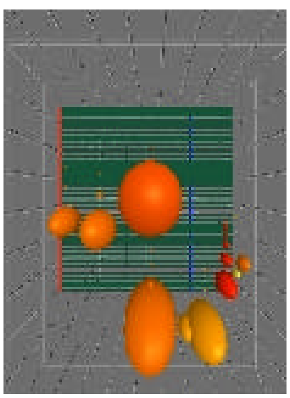
The performance depends greatly on the sophistication of the sounds - that is, on the number of partials per sound and the number of active controls for each partial. The speed increases with the number of nodes, but load imbalances result in a sublinear speedup. Nevertheless, in most cases, this part of the code performs in real time. A contemplated improvement is to eliminate the mixer, which collects the output from the other processors and constitutes a potential bottleneck. In a future version, each node would write directly to the output file as it computes the samples.

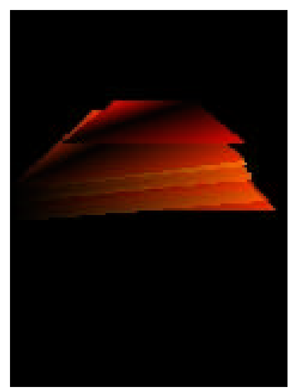
2.5. The CAVE, ImmersaDesk, and M4CAVE

"A.N.L.-folds" could also be considered a multimedia work consisting of computer-generated tapes and images projected in the Cave Automatic Virtual Environment (CAVE) or on the ImmersaDesk. The CAVE is a room-size (10X10X10 ft.), multiperson, high-resolution 3D video and audio environment, where stereo images are projected on three walls and the floor. The ImmersaDesk is a 2D scaled-down version of the CAVE.

The visual images projected in the CAVE are generated directly from DIASS score files with the help of M4CAVE, a program written in C++ and OpenGL. The images are computed on the fly and are made to correspond exactly to the sounds through a one-to-one mapping between visual attributes and sound qualities. The user can interact by turning on and off any of the parameters and by determining the way in which sounds are represented.

Our experiments with music visualization consist so far of three alternative ways to graphically represent the sounds. In the first one, partials are represented as spheres whose size is proportional to loudness and whose height corresponds to frequency. An optional grid in the background shows the octaves divided into twelve equal increments. A sound's position in the stereo field determines the placement of the sphere between left and right in the room. Visual objects rotate or pulse when tremolo or vibrato is applied, and their color varies when reverberation is present. The second representation consists of a cloud of confetti-like objects that react to sound attributes: loud sounds produce more agitation in the cloud, its position in the room is influenced by the panning, and frequency and reverberation change the color of the objects. In the third representation, a collection of colored planes float in space and change position, color, and degree of activity according to the same sound parameters. All three representations are an accurate rendition into images of musical data. The basic difference between them is that the first one, the spheres, maps more sound qualities into the images than do the other two.





Recent work on M4CAVE involved the creation of a new menu that allows the viewer to modify each representation's mappings and to switch between graphic representations during the rendition of the same file.

3. Sonification
Since sounds and images can be created from the same data (score file), DIASS and M4CAVE have become tools for sonification.

We subscribe to the idea that sonification is the faithful rendition of data in aural images (sounds). The aural images should reflect the details of the data in an impartial manner; they should definitely not give a limited interpretation or caricature. A good sonification tool is like a neutral and honest broker, who leaves the task of filtering and finding meaning to the user.

This suggests the need for a tool that is both sufficiently powerful to handle data sets with a large number of degrees of freedom and sufficiently flexible to be responsive to subtle fluctuations in the data. DIASS is such a tool because it can handle a relatively large number of degrees of freedom and offers precise and detailed control over each of them. Behind DIASS there is the view that the world of sound is a multidimensional space, whose various coordinates (or "sound parameters," to use the terminology introduced by Xenakis and the integral serialists in the 1950s) correspond to perceived qualities of sounds and that, in turn, constitutes a solid basis for possible isomorphisms between the aural domain and other multidimensional spaces.

3.1. Sonification experiments

Two series of sonification experiments have been produced so far. The first one originated with a computational chemist who studied the binding of a carbon atom to a protonated thiophene molecule. The data represented the difference in the energy levels before and after the binding at 128X128X128 points regularly arranged in space on the computational mesh points in a cubic lattice. We also looked at a simpler case, that of a molecular structure represented in a basis of Legendre polynomials.

Play Sound File

In both instances we were dealing with static data, and we identified time with a spatial coordinate. We arbitrarily selected one of the coordinate axes of a rectangular (xyz) coordinate system (the x axis, say) for this purpose and sonified the data in planes parallel to this axis. The time to traverse a plane over its full length was kept fixed at 30 seconds. Amplitude and frequency were used alternatively to represent the changes in the data from point to point on this 128X128 grid.

The second experiment involved data from a numerical simulation in material science. The scientists were interested in patterns of motion of magnetic flux vortices through a superconducting medium represented by 384X256 points in a rectangular plane. As the vortices are driven across the medium, from left to right, by an external force, they repel each other but are attracted by defects in the material. In this experiment frequency was used to represent the movement of vortices in the plane and changes in loudness were connected to changes in the speed of a vortex. A traveling window of constant width was used to capture the motion of a number of vortices simultaneously.

The results of the investigation, which is ongoing, have met our expectations so far: sounds were produced that conveyed information about the qualitative nature of the data. DIASS proved to be a flexible and sophisticated tool capable of rendering subtle variations in the data. At the same time, these experiments showed that DIASS, at least in present form, has its limitations.

3.2. DIASS as an instrument for sonification

One limitation of DIASS concerns the sheer volume of data in scientific sonification. While the composition of a musical piece (the original intent behind DIASS) typically entails the handling of a few thousands sounds, there are over 2 million data points in the chemistry problem, a difference of several orders of magnitude. By the same token, while a typical amplitude envelope for a partial or sound involves ten or even fewer segments, both experiments required envelopes with well over 100 such segments.

Another difficulty encountered was the fact that both experiments required sounds to be accurately located in space. While the usual panning used in music is very effective in pinpointing the source on a horizontal line, suggesting the height of a sound is a major challenge. We hope that additions to the software as well as a contemplated eight-speaker system in the CAVE will help us getting closer to a realistic 3D representation of sounds.

Finally, in order to become an effective sonification tool, DIASS requires real-time capabilities.

All three concerns are being addressed in the new C++ version of DIASS.

4. Larger issues

The fact that a single tool - a digital synthesis instrument - can be useful both in scientific research and in the creation of musical works has implications that touch on the relationship between science and art, their parallel or divergent goals, and ways of informing us about the world.

The project is not only of practical interest, for music composition as well as scientific research. It also has more fundamental, theoretical aspects pertaining to a rigorous mathematical description of sound events and to the way they are perceived, alone or in conjunction with visual images. Besides bringing in specialized knowledge, the multidisciplinary approach also offers an integrated view of these broader issues at a higher intellectual level. It also showed us that the way in which both composers and scientists perform their tasks might be about to change in several respects.

For the composer, the use of cutting-edge technology not only means the possibility to experiment more and obtain results faster, but also points to the need to re-evaluate basic building blocks and aesthetic habits. How do we define in a rigorous way a sound or a composition? Is the result of a sonification experiment less "musical" than, say, the thunder in the Alpine Symphony or Respighi's chirping birds? Composing the internal makeup of sounds is as important for some of us as putting them together in a piece. But no single person can possibly have the required expertise in so many and diverse realms: music, psychoacoustics, mathematics, and computer science. The concept of a manifold composition - impossible to realize without fast computers - challenges the very idea of an "art object" fixed in its appearance, available for repeated inspection or used as an investment.

If sonification of complex data sets proves to be as useful as we indeed believe it is, the computational science community will face some challenges too. Ours is a visually oriented culture, we watch rather than listen. With an open mind and an awareness for unusual sonorities, scientists may discover the benefits of the world of sound and appreciate its use in significantly increasing the bandwidth of the human/computer interface.

Science and art are much closer than we were told: they both experiment and strive to present through their specific means a view of reality.

References

- J. Beauchamp - *Music 4C Introduction*, Computer Music Project, School of Music, University of Illinois, Urbana IL, 1993. WWW:htp://cmp-rs.music.uiuc.edu/cmp/software/m4c.html.
- T. DeLio - *Circumscribing the Open Universe*, University Press of America, Lanham MD, 1984.
- U. Eco - *The Role of the Reader*, Indiana University Press, Bloomington IN, 1979.
- *Interface*: Journal of New Music Research vol. 16 (1987). Issue dedicated to the topic "Open Structure in Twentieth-century Music".
- H. G. Kaper, D. Ralley, J. M. Restrepo, and S. Tipei, - *Additive Synthesis with DIASS_M4C on Argonne national Laboratory's IBM POWERparallel System (SP),* Proc. 1995 Int'l Computer Music Conference (Banff, Canada), pp. 351-352.
- H. G. Kaper, D. Ralley, and S. Tipei - *Perceived Equal Loudness of Complex Tones: a Software Implementation for Computer Music Composition,* Proc. 1996 Int'l Conference in Music Perception and Cognition (Montreal, Canada), pp. 127-132.
- H. G. Kaper, S. Tipei, and E. Wiebel - *High-performance Computing, Music Composition, and the Sonification of Scientific Data,* (1997). Preprint ANL/MCS-P690-0997, Matematics and Computer Science Division, Argonne National Laboratory.
- C. Kriese, and S. Tipei - *A Compositional Approach to Additive Synthesis on Supercomputers*, Proc. 1992 Int'l. Computer Music Conference (San Jose, California), pp. 394-395.
- H. Pousseur - *La nova sensibilita musicale*, Incontri Musicali vol. 2, 1958, p. 25.
- K. Stockhausen - *...how time passes...*, Die Reihe vol.3, Universal Edition, 1959, p. 32.
- S. Tipei - *The Computer: a Composer's Collaborator*, Leonardo vol. 22 no. 2, 1989, pp. 189-195.
- S. Tipei - *Manifold Compositions - A (Super)computer-assisted Composition Experiment in Progress*, 1989 Int'l Computer Music Conference (Columbus, Ohio), pp. 324-327.
- S. Tipei - *A.N.L.-folds. mani 1943-0000; mani 1985r-2101; mani 1943r-0101; mani 1996m-1001; mani 1996t- 2001(1996).* Preprint ANL/MCS-P679-0897, Mathematics and Computer Science Division, Argonne NationalLaboratory.