Developing Music Notation for the Live Performance of Electronic Music

Alan Tormey
Ph.D Princeton University
Associate Creative Director, Eclectic Laboratory Chamber Orchestra

Preamble

This paper arises out of my time performing with the Princeton Laptop Orchestra (PLORK), an ensemble comprising 15 to 20 laptops, over 90 loudspeakers, and a variety of performance interfaces. After being with PLORK for three seasons, dozens of performances, and approximately a score[1] of individual pieces, I noticed several general trends – most notably that a preponderance of the works composed for PLORK rely heavily upon some form or another of improvisation, aleatory, and/or open form composition. While these are, of course, all valid and often effective means of creation, it is my intuition that, in certain instances, the choice to use these methods was made, in large part, as a reaction to the overwhelming novelty of the ensemble and the lack of any defined and effective paradigms within which to develop and communicate more strictly specific musical ideas.[2] They did not know what they wanted to do or else did not know how to communicate what they wanted to do and so were left with the negative solution of entrusting the ensemble to do it in their stead. Even in pieces where the composers had a very clear and well-defined vision of the performance material, the works were often, if not usually, taught by rote – ceasing to exist if and when there is enough turnover of significant orchestra personnel. Regardless of the irony inherent in the bleeding-edge of Western music depending on pre-historic methods of transmission, it is, the author believes, something of a tragedy that this music is generally not available to be learned and performed by the many new laptop-centric ensembles forming around the world.

This paper hopes to point towards the beginnings of certain solutions for some of these problems. After identifying basic issues in notation that are germane to the difficulties of notating music for digital instruments, the principles behind hypothetical notational solutions for three types of gestural interface instruments are looked at in

---

[1] pun intended

[2] A similar, but distinct, predicament is related in Peters, Nils, Georgios Marentakis, and Stephen McAdams "Current Technologies and Compositional Practices for Spatialization: A Qualitative and Quantitative Analysis" *Computer Music Journal* 35.1 (Spring 2011): 17.

closer detail. Although it ostensibly deals solely with music notation the paper is also, in many ways, asking to what extent the closed work will be a part of large ensemble music for laptops and/or gestural interface instruments.

Instrument Neutrality

Traditional music notation is extraordinary, both for the number of functions it is able to simultaneously fulfill and for the efficiency with which it fulfills them. One interesting feature of traditional notation is that it is, more or less, indifferent to the instrument realizing it. If we visualize the melody "Twinkle, Twinkle Little Star," for example, it will sound identifiably the same on just about any instrument of Western pitch. However, depending on whether the player is using a flute, violin, trumpet, or some other instrument, that same melodic notation will elicit a very different set of physical actions from the performer. Part of the reason for this is that traditional music notation primarily works backwards from effect to cause by representing a desired sonic output that the performer reverse engineers into the physical actions that are required in order to produce the music. Seen through this lens, the practice of learning a traditional Western instrument becomes the act of internalizing a set of algorithms that will effect a one-to-one correlation between the physical gestures put into the instrument and the sound that will come out of it. Because of these hypothetical algorithms, standard notation can be understood to simultaneously represent both sounds and gestures.

Geometry of Physical Action in Traditional Music Notation

Music notation's linkage of sound and action is an intuitive one since our understanding of pitch is generally spatial in character.[3] [4] Not only do we conceive of pitch in spatial terms like *high* and *low*, but the actual physics of a note are spatial too, since the frequency an instrument plays is a direct function of the length and thickness of its string or tube. Although different instruments have different relationships to physical space, strings, woodwinds, and keyboards each (basically) control their pitch along a one-dimensional axis (or, in a case such as the violin, we can imagine a series of parallel one-dimensional pitch axes). The musical staff literally represents this one-dimensional pitch

---

[3] Zbikowski, Lawrence M. "Metaphor and Music Theory: Reflections from Cognitive Science" *Music Theory Online,* 4.1 (1998): Web.

[4] Shakila, Shayan, Ozge Ozturk, and Mark A. Sicoli. "The Thickness of Pitch" *Sense & Society,* 6.1 (2011): 96-105.

geometry along the Y-axis. This leaves the X-axis free to metaphorically represent time as a perpendicular one-dimensional space – an important feature of standard notation that we will return to.

Problems of Digital Notation/Gestural Interfaces Instruments

A gestural interface instrument is a system in which the motions/actions of a performer are captured by hardware and interpreted by software to create musical sound. When working with these instruments, the composer not only has to answer the question of what the music will be, but must now also answer the question of how will the music actually be performed. Traditional instruments are well understood in terms of the relationship between gesture and sound, but no such practice exists for these new digital devices. Each new piece of music requires its own piece of custom software that will create a set of relationships between the physical movement of a performer and the sounds that those motions will ultimately create – sounds that will likely be different from piece to piece. This decoupling of gesture from sound radically changes not only the tasks that music notation must fulfill but also the ways in which it can fulfill them. **WITHOUT A STABLE SOUND/GESTURE CORRELATION, A NOTATION SYSTEM FOR GESTURAL INTERFACES MUST SIGNIFY THE PHYSICAL INPUT RATHER THAN THE SONIC OUTPUT IF IT WISHES TO EFFECTIVELY COMMUNICATE SPECIFIC PERFORMANCE INSTRUCTIONS TO AN INSTRUMENTALIST.** Likewise, since different instruments each have different gestural practices, new notation systems will have to be developed for each individual instrument or, at the very least, each general family of instruments that share a basic gestural practice. The author holds it as axiomatic that these systems can, ideally, share the largest possible number of common principles, both with each other and with traditional notation – so as to take advantage of the latter's large pre-installed user base.

Discrete events vs. continuous data; Virtual Handbells

In the study of Human-Computer Interaction, one of the first distinctions made is between interfaces that measure discrete events and interfaces that measure continuous changes in data.[5] Acoustic instruments and their notation systems primarily deal in discrete events – individual pitches and rhythms are put forth in succession. **THEREFORE,**

[5] Verplank, Bill. "Interaction Design for Human Use" Lecture Sept. 1998. Accessed at <http://hci.sapp.org/lectures/verplank>

DIGITAL INSTRUMENTS THAT OPERATE PRIMARILY BY TRANSMITTING DISCRETE POINTS OF DATA SHOULD BE THE EASIEST TO ACCOMMODATE TO THE PRACTICES OF TRADITIONAL NOTATION.

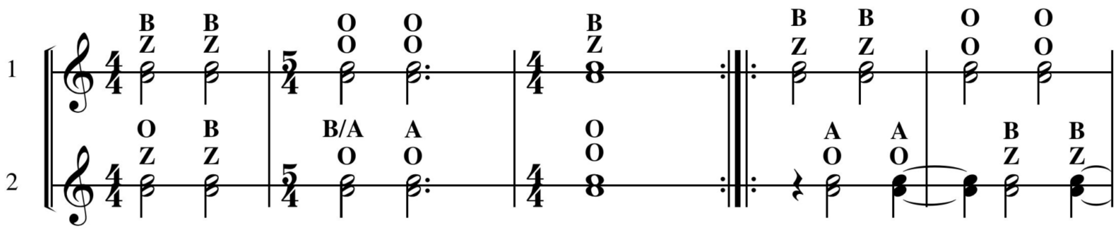
EXAMPLE 1 -AUDIO

Example 1 is a passage of music composed for a duo of virtual handbell instruments that each use a Nintendo Wii and Nunchuk controller system as their physical interface.[6] The instrument works by triggering a MIDI note-on message whenever the controller's motion changes from down to up with a certain amount of force. The note number triggered can be changed based on which of the controllers' four main buttons are engaged. Because the virtual handbells articulate a series of discrete events rather than continuous data, its behavior is not dissimilar from an acoustic instrument. Therefore, all three of the example notations can retain traditional notation's spatialization of time along the X-axis.

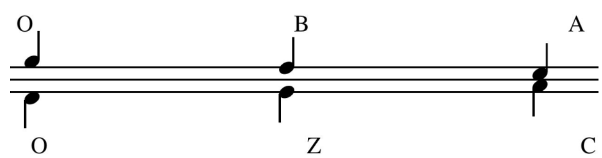
Examples 2 – 4 are variant notations for this music. Ex. 2 shows the passage in standard notation, while Ex. 3 demonstrates a simple notation that places rhythms and controller buttons on parallel X-axes. While this is a clear, easy-to-read notation it does little, if anything, to communicate any musical aspects of the passage aside from its rhythm. On the other hand, Example 4 is something of a hybrid between the previous two notations, combining the simplicity of Example 3 with the added depth afforded by a multi-dimensional system such as standard notation. As the key shows, each button is assigned a different place on the staff. Like Ex. 3, Ex. 4 illustrates performative input rather than sonic output but, like Ex. 2, it also utilizes vertical information. While the Y-axis no longer necessarily correlates to register (although the software could certainly be programmed that way), the buttons' placement on the staff makes it easier to see the shapes of performance that exist within the part. Examining the part at player 2, measure 1, the standard notation reads as an idiosyncratic, angular moment comprising cross-voicing, compound intervals, and large registral leaps. In contrast, Ex. 4's repeated and step-wise noteheads reflect the actual simplicity of the performance. It looks like it plays.

[6] Tormey, Alan. *Etude no. 2 (music for virtual handbells)* 2009.

EXAMPLE 2 – STANDARD NOTATION



EXAMPLE 3 – SIMPLE BUTTON NOTATION



EXAMPLE 4A – HYBRID NOTATION KEY

EXAMPLE 4B – HYBRID NOTATION

Despite being a possible and, in some ways, even effective representation of the passage, standard notation is not the ideal solution. Furthermore, beyond the problems described in the previous paragraph, a single click of the mouse will render it completely useless. Example 5 is the exact same set of performance actions following a change of synthesizer preset. Since the computer has now drastically altered the sonic output created by performance actions, the version of the passage written in standard notation is now an inaccurate and misleading representation of the music's sound. While we could transcribe the new pitch content and alter the standard notation to reflect the new synth patch, doing so creates needless and confusing work for the performer who would need to learn two different performance keys in order to repeat the same set of phsyical actions. In imagining a situation where Examples 1 and 5 would both coexist in the same piece, we can see that the score will be much more effective when *prescribing* the performer's physical input rather than *describing* the music's sonic output.

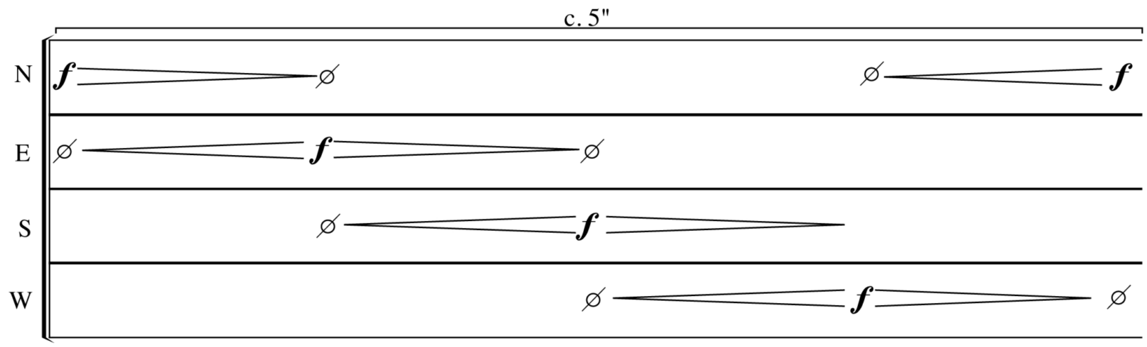
EXAMPLE 5 - AUDIO

Musical Dynamics and Continuous Control

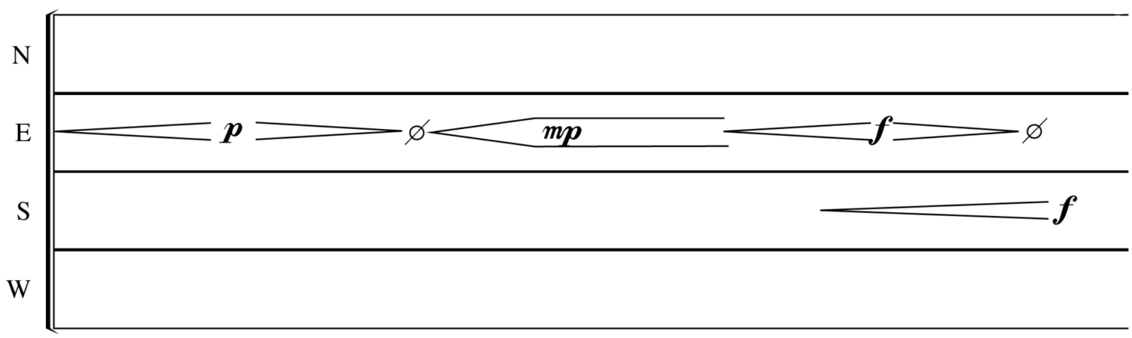
Thinking back to standard notation, we recall that pitch and rhythm are notated as a series of discrete values on the Cartesian grid. The other main elements of music – tone color and dynamics have, historically, been indicated as single global parameters if and when they were notated at all. Over time, however, the notation of dynamics has developed into one instance where the traditional system accommodates the representation of continuous data. Like pitch and time, volume is also understood as a linear one-dimensional value, which is located on a second X-axis parallel to the staff.

Simple Continuous Control: Joystick Instruments

The use of dynamics as a model for the representation of continuous musical data inspired the following experimental notation system for joysticks. The joystick is constructed around four sensors, one for each of the four cardinal directions. Diagonal motions register by activating two sensors at once. The system illustrated in Exx. 6 and 7 maps the linear one-dimensional space of musical dynamics – from *niente* to *forte* – onto each of the joystick's four cardinal directions: North, East, South, and West (*forte* equals the outermost distance in any direction). Crescendo and decrescendo indications drive motion between joystick positions. Example 6 indicates a clockwise circle around the outer circumference of the joystick, while Example 7 indicates a swell and shorter turn. Notice that, by treating leftward motion as a separate parameter from rightward motion and **not** its opposite, leftward motion can be represented while still retaining a left-to-right spatialization of time.



EXAMPLE 6 – JOYSTICK NOTATION, CLOCKWISE CIRCLE (BEGINS WITH STICK AT FULL NORTH)

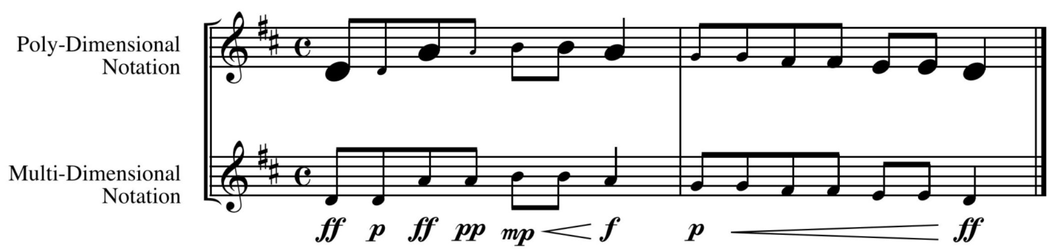


EXAMPLE 7 – JOYSTICK NOTATION, SWELL AND TURN (BEGINS WITH STICK AT CENTER)

Poly-Dimensional Notation

While the joystick notation does offer a successful model for dealing with continuous data, the joystick's physical mechanics limit its spatial capabilities to contiguous points on the two-dimensional Cartesian plane. However, other new interfaces can interpret fuller, more detailed ranges of motion and gesture. A Wacom tablet, for example, measures Cartesian position and the pressure, altitude, and azimuth of the stylus. Although effective for its task, the joystick notation is really nothing more than four one-dimensional values lain in parallel and is not up to the added complexity that comes with added dimensions of data. As we move into interfaces that can interpret more varied ranges of motion, clearer and more efficient methods are called for.

Given the need to not only represent multiple streams of information, but to do so in a way that can be efficiently interpreted, a poly-dimensional notation that can represent multiple streams of information via a single indication would, in many cases, help simplify the transmission of gesture-based musical scores. Returning again to dynamics – the most familiar and fully realized system for describing continuous data in acoustic notation – there is a relatively obscure mid-twentieth century innovation that uses notehead size to fuse the multiple dimensions of pitch and dynamics into a single poly-dimensional indication.[7]



EXAMPLE 8 – POLY-DIMENSIONAL NOTATION

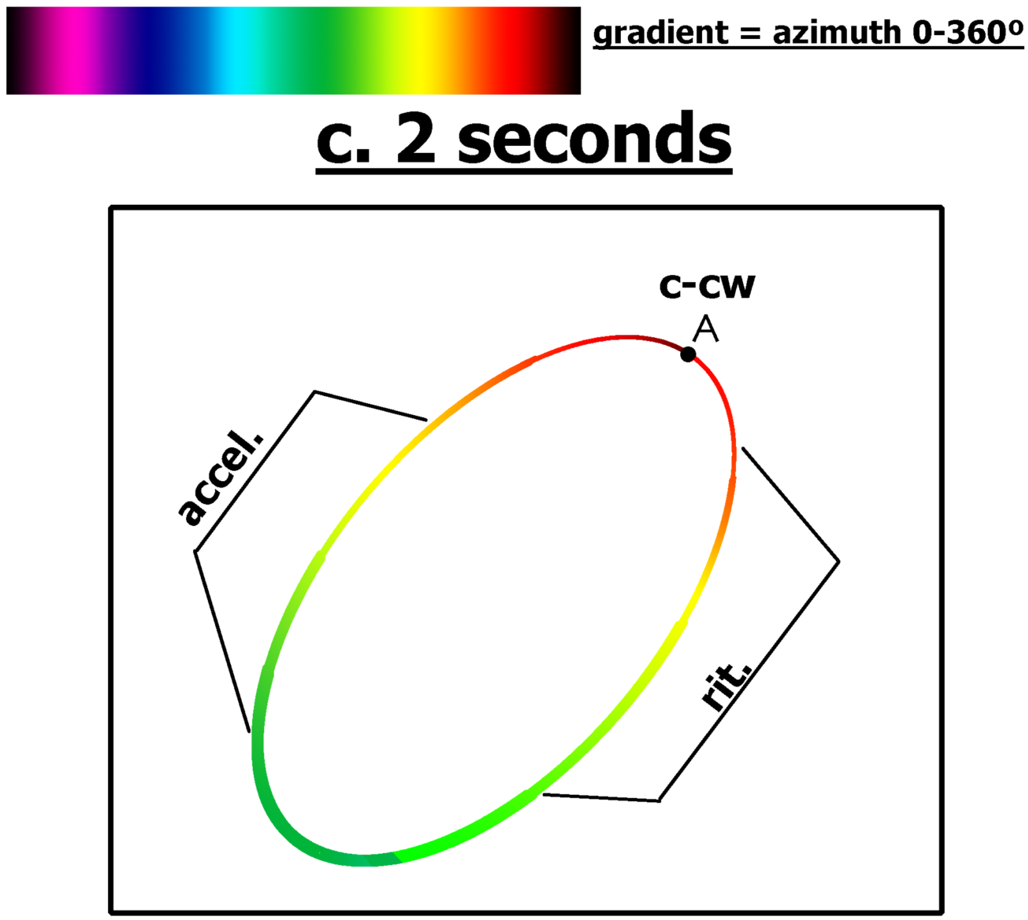
In Example 9 – a potential pen and tablet notation – a line traces the path of the pen starting from a dot. The thickness of the line indicates relative pressure, while color is used to indicate the pen's azimuth as it travels along its path. The pen's altitude measurement could easily be added to this by modifying some part of the azimuth color, such as the saturation level. As Example 10 shows, this notation can be easily accommodated to an ensemble context by utilizing proportional and cue-based notational practices common to much mid-century art music.

Video and Time

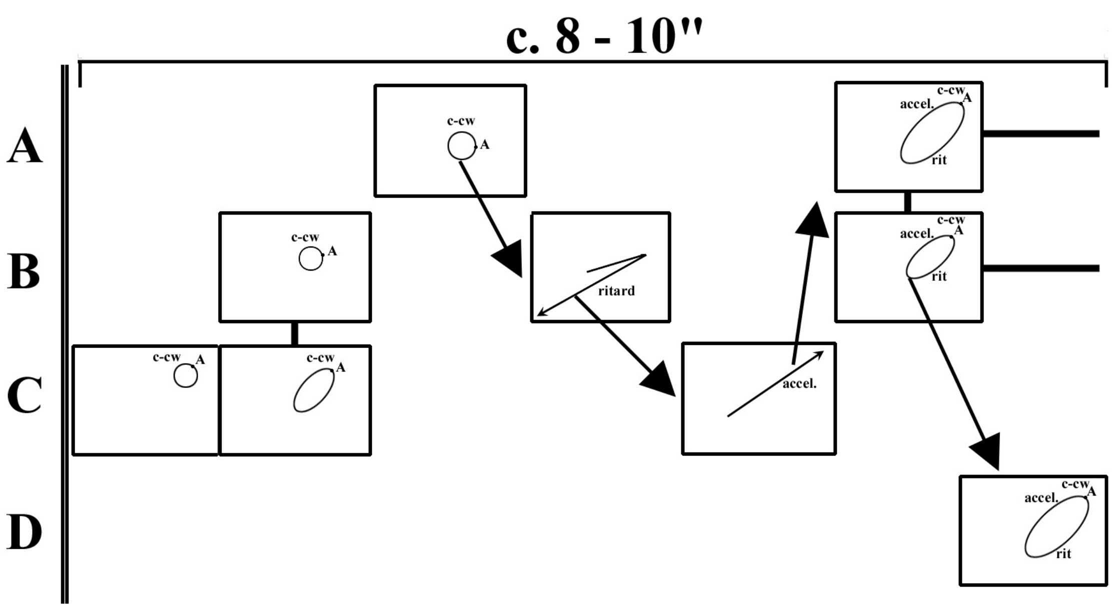
Obviously, this notation's weakest element is its relative imprecision and inelegance regarding time – a seemingly ever-present stumbling block in the path of clear, efficient gestural notation. But what if time wasn't notated as a static function of space, but remained fluid and temporal? What if time was represented as time? Devices such as the Wiimote and the Wacom tablet were designed for the needs of the gaming and computer industries, and implicitly assume that video monitoring, not print, will be the primary medium of reception. Using video as a notation medium frees time from having to be represented spatially. Video game designers have already offered solutions for video-based notation systems. Games such as *Guitar Hero* implement an innovative new system of time-based notation that has carried on in recent musical applications by companies such as Smule.[8]

---

[7] Stone, Kurt. *Music Notation in the Twentieth Century* New York: Norton, 1980.
[8] See, for example, demonstration videos of Smule's *Leaf Trombone*.



Example 9 - Pen and Tablet Notation



Example 10 – Uncolored Pen and Tablet Notation in an Ensemble Context

Using video as the new medium of notation makes intuitive sense – why shouldn't music made on computers utilize computers for its scoring and transmission? Virtually all the musicians who would make and/or perform this music have access to the basic necessary technology and, in many, if not most, cases, a computer will already be on stage in order to run the gestural devices.

So, while the material burdens of video notation may not be great, there are several practical problems raised by a paradigmatic shift to video. What is the purpose of a video score? The static representation of time is fundamental to the traditional score's ability to function doubly as both a set of real-time instructions that will instantiate a musical performance and also as an atemporal representation of the composition that can be used for analysis and overview. A **NOTATION THAT REPRESENTS PHYSICAL INPUT RATHER THAN SONIC OUTPUT IS UNLIKELY TO BE ABLE TO RETAIN ITS BREADTH OF FUNCTIONALITY, TO BE SIMULTANEOUSLY PRESCRIPTIVE AND DESCRIPTIVE, TEMPORAL AND ATEMPORAL, PERFORMATIVE AND ANALYTIC.** Composers will have to either expend the resources to employ a video artist or else learn animation and/or new computer languages – a proposition that would necessitate substantial, controversial, and expensive changes to music pedagogy. What, exactly, are the implications of asserting that animation is now a prerequisite skill in the study of music? On the performance side, a time-based system also brings inherent difficulties related to practicing the instrumental part, which may lead towards a trend of composing overly simplistic material in reaction to performers' restricted ability to rehearse particular spots and passages.

Video notation will certainly not solve all of the problems brought about by digital instruments' incompatibility with traditional notation. In the future, composers will not only have to choose how to notate their scores, but decide what function they want their score to ultimately serve. Rather than quest after a single system/single artifact solution that continues to combine the many functions of traditionally notated scores, composers should encourage themselves to simultaneously employ multiple systems of communication as they become necessary to the needs of the work.