



This is to certify that the

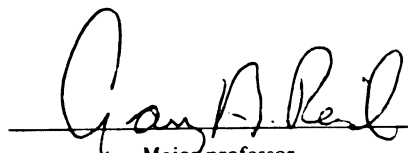
thesis entitled
 AN INVESTIGATION OF THE CAPACITY OF THE
 MUSICAL INSTRUMENT DIGITAL INTERFACE TO RENDER
 MUSICAL RHYTHM AND THE SENSITIVITY OF AN
 AUDIENCE TO MACHINE-PRODUCED MUSICAL RHYTHMS

presented by

Kenneth James Tanner

has been accepted towards fulfillment
 of the requirements for

M. A. degree in TELECOMMUNICATION


 Major professor

Date 6/2/92

LIBRARY
Michigan State
University

PLACE IN RETURN BOX to remove this checkout from your record.
TO AVOID FINES return on or before date due.

DATE DUE	DATE DUE	DATE DUE
1 2 1 1 0 3 OCT 0 8 2002	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____
_____	_____	_____

MSU Is An Affirmative Action/Equal Opportunity Institution

c:\clrc\datedue.pm3-p.1

AN INVESTIGATION OF THE CAPACITY OF THE MUSICAL INSTRUMENT
DIGITAL INTERFACE TO RENDER MUSICAL RHYTHM AND THE SENSIVITY
OF AN AUDIENCE TO MACHINE-PRODUCED MUSICAL RHYTHMS

By

Kenneth James Tanner

A THESIS

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

MASTER OF ARTS

Department of Telecommunication

1992

ABSTRACT

AN INVESTIGATION OF THE CAPACITY OF THE MUSICAL INSTRUMENT DIGITAL INTERFACE TO RENDER MUSICAL RHYTHM AND THE SENSITIVITY OF AN AUDIENCE TO MACHINE-PRODUCED MUSICAL RHYTHMS

By

Kenneth James Tanner

This thesis investigates the capacity of the Musical Instrument Digital Interface (MIDI) to produce human-like musical rhythms. Human musical rhythms were collected and analyzed using MIDI-equipped drum sets, a computer-based MIDI sequencer, and two drummers. Twenty performances were statistically analyzed to determine how they differed from the mechanical performance model. Significant differences were found between how a human drummer performed a given rhythm and how a MIDI sequencer or drum machine would normally produce it. Pairs of human and machine-produced rhythms, equal in all respects but individual note durations, were presented to an audience to determine listener preference for human versus machine-produced rhythms. There was no majority preference overall for either human or machine-produced renditions, suggesting duration differences, such as those found in the test stimuli, do not alone strongly bias audience preference for rhythmic performances. Advice is given on how to edit a mechanical model drum pattern so it more closely resembles a human performance.

Copyright by

Kenneth James Tanner

1992

ACKNOWLEDGMENTS

Thanks to Gary Reid for his kindness, patience and guidance. Also, for his musical insights and technical expertise, and for introducing me to this challenging problem.

Thanks to Dr. Carry Heeter, a member of my committee, especially for her help with the audience test questionnaire. The audience test was definitely more successful because of it. Thanks also for the thoughtful feedback on the results.

Thanks to committee member, Bob Albers for his time, interest and geniality. It has been a pleasure to meet and talk with him on various subjects over the past three years.

Thanks to my old friend and band mate Chris Moore, who served as a drummer in one of the studio experiments. Thanks to Scott Kuizema for drumming in the other studio experiment and to Larry Kuizema for helping to set it up.

Thank you to my father, William Tanner, for the use of his computer facilities and for his help in the mechanics of formatting and printing the paper.

Thanks, finally, to my wife, Tammie, for her help, encouragement and patience throughout this three year adventure.

Ken Tanner

May 1992

Table of Contents

LIST OF TABLES	viii
LIST OF FIGURES	xii
I. INTRODUCTION	1
Origin of MIDI	1
What is MIDI?	3
Synchronization and Timing	5
System Real Time Messages	6
Recording and Playing Back on a MIDI Sequencer	8
MIDI Delay	10
II. MIDI AND RHYTHM RESEARCH	15
Musicology Research on Rhythm	16
Specific Studies	17
III. IMPROVING SEQUENCED RHYTHM QUALITY	23
IV. TAPE AND PENCIL REGISTRATION OF MUSICAL RHYTHM	31
Method	31
Results	32
Discussion	39
V. MIDI SEQUENCER REGISTRATION OF MUSICAL RHYTHM	42
Pretest of MIDI Sequencer Registration Method	46
Results of Studio Session One	53
Pattern 1	54
Pattern 2	60

Pattern 3	72
Results of Studio Session Two	79
Pattern 4	79
Pattern 5	82
Pattern 6	87
Discussion of Studio Experiments	90
VI. AUDIENCE TEST	94
Method	94
Results	95
Discussion	99
VII. RECOMENDATIONS FOR MUSICIANS AND FUTURE RESEARCHERS	105
For Musicians	105
For Future Research	109
APPENDICES	
Appendix A - Sample Audience Questionnaire . . .	111
LIST OF REFERENCES	116
GENERAL REFERENCES	118

LIST OF TABLES

Table Number	Title	Page
1	Tape and Pencil: Correlations, Actual and Rational-Mechanical Durations.	37
2	Tape and Pencil: Slopes and Intercepts.	37
3	Tape and Pencil: Actual and Rational-Mechanical Durations for the Three Tape and Pencil Registrations.	39
4	Tape and Pencil: T-Tests of the Three Eighth Notes.	47
5	Pretest: Rotated Factor Loadings for the 10 Versions.	48
6	Pretest: Correlations, Actual and Rational-Mechanical Durations.	49
7	Pretest: Slopes and Intercepts.	49
8	Pretest: Actual and Rational-Mechanical Total Durations.	50
9	Pretest: Differences From Rational-Mechanical Total Duration.	52
10	Pattern 1: T-Test of Notes 2 and 3.	54
11	Pattern 1: T-Test of Notes 4 and 5.	54
12	Pattern 1: T-Test of Notes 5 and 1.	55
13	Pattern 1: T-Test of Velocity for Notes 1 and 4.	55
14	Pattern 1: Number of Consequential Differences.	56
15	Pattern 1: Slopes and Intercepts.	57

Table Number	Title	Page
16	Average Tempos for Pattern 1, Versions One Through Five.	57
17	Pattern 1, Version 1: Comparison of the Accuracy of Linear and Quadratic Models in Macro tempo Prediction.	58
18	Pattern 2: T-Tests of Consecutive Notes with Equal Rational-Mechanical Durations.	61
19	Pattern 2: T-Tests of Beat Durations, Drums.	62
20	Pattern 2: T-Tests of High-Hat Note Durations.	62
21	Pattern 2: Correlations Between Drum and High-Hat Durations.	63
22	Durations for the First Two Measures of Pattern 2, Version 4.	64
23	Pattern 2, Version 1: Drum/High-Hat Offsets.	65
24	Pattern 2: Drum/High-Hat Offsets, Bass and Snare Drum.	66
25	Pattern 2: Drum/High-Hat Offsets, Bass Drum Only.	67
26	Pattern 2: Drum/High-Hat Offsets, Snare Drum Only.	67
27	Pattern 2: Offset Summary Statistics for Overall, Bass Drum and Snare Drum Distributions.	68
28	Pattern 2: Proportion of Time Drum Hits Before, After, or Simultaneous with High-hat.	68
29	Pattern 2: Number of Consequential Differences.	69
30	Pattern 2: Slopes and Intercepts.	69

Table Number	Title	Page
31	Pattern 2: Average Tempos for the Five Versions.	70
32	Pattern 3: T-Test of Notes 2 and 3.	73
33	Pattern 3, Measure B: Consecutive Note T-Tests	73
34	Pattern 3: T-Test to Gauge Performance of Rest in Measure B.	74
35	Pattern 3: Offset Summary Statistics for Overall, Bass Drum and Snare Drum Distributions.	75
36	Pattern 3: Proportion of Time Drum Hits Before, After, or Simultaneous with High-Hat.	75
37	Pattern 3: Number of Consequential Differences.	76
38	Pattern 3: Average Tempos.	76
39	Pattern 3: Slopes and Intercepts.	77
40	Pattern 4: Summary Statistics for Overall, Bass and Snare Drum Distributions.	80
41	Pattern 4: Percentage of Simultaneous and Non-Simultaneous Hits by Drum.	80
42	Pattern 4: Correlations Comparing Pattern 1, 2 and 4 Residual Scores.	82
43	Pattern 5, Version 1: Summary Statistics for the Overall, Bass and Snare Drum Distributions.	83
44	Pattern 5, Version 1: Proportion of Time Drum Hits Before, After or Simultaneous with High-Hat.	83
45	Pattern 5: Slopes and Intercepts.	84
46	Pattern 6: Offset Summary Statistics for the Overall, Bass and Snare Drum Distributions.	87

Table Number	Title	Page
47	Pattern 6: Percentage of Simultaneous and Non-Simultaneous Hits.	88
48	Pattern 6: Slopes and Intercepts.	88
49	Overall Macro tempo Behavior.	92

LIST OF FIGURES

Figure Number	Title	Page
1	Twenty-Four PPQ Quarter-Note Subdivided	6
2	Timing Clock Message Inserted Between a Status and Data Byte	6
3	Sample Swing Point Setting for the Roland R-8 Human Rhythm Composer	24
4	Sample Pattern with 1:1 High-Hat Duration Ratio	27
5	Sample Pattern with 13 Percent Skew in High-Hat Durations	28
6	Using Tied Quintuplets to Represent a 3:2 Relationship	28
7	Approximating a 3:2 Relationship with Dotted Eighth-Sixteenth Combinations	29
8	Micro Tempo Map for "I'm Ready"	33
9	Micro Tempo Map for "Fool and Me"	34
10	Micro Tempo Map for "Good Lovin' Gone Bad"	34
11	Residual Map for "Good Lovin' Gone Bad"	38
12	Pretest Drum Pattern Notation	47
13	Micro Tempo Maps for Pretest	51
14	Notation for Pattern 1	54
15	Residual Maps for Pattern 1	59
16	Notation for Pattern 2	60
17	Residual Maps for Pattern 2	71
18	Notation for Pattern 3	72
19	Residual Maps for Pattern 3	78

Figure Number	Title	Page
20	Notation for Pattern 4	79
21	Residual Map for Pattern 4	81
22	Notation for Pattern 5	82
23	Pattern 5 Micro Tempo Maps	85
24	Residual Maps for Pattern 5	86
25	Notation for Pattern 6	87
26	Residual Maps for Pattern 6	89
27	10 Msec Note Recorded at Tempos of 100 and 60	108

INTRODUCTION

Many people have seen or heard of the acronym, MIDI, which stands for Musical Instrument Digital Interface, and most have heard music produced with this technology. The ubiquitous drum machine, churning out its precision rhythm accompanying a solo performer at a local bar, is a MIDI device. One hears this mechanical drummer on many stations up and down the radio dial, or, virtually store to store in the shopping mall. Less conspicuously, MIDI directs synthesized orchestras and ensembles in the sound tracks of videos, movies, television shows and commercials, often in combination with non-synthesized instruments. Using MIDI, a composer can hear a realization of his score without the aid of other musicians. Complex tasks in the recording studio can be automated, providing a new measure of efficiency and creative freedom.

While MIDI has become a valuable tool in media production, it is also a promising aid in music research. The ability of MIDI to precisely manipulate certain music parameters (e.g., duration and amplitude of notes) can prove useful in experiments on music performance and audience perception.

This investigation will test how well MIDI is able to represent certain musical rhythms, and how sensitive an audience is to MIDI synthesis of those rhythms. In doing so, it will draw on the existing body of musicology research on musical rhythm. Finally, it hopes to formulate some rules to guide the musician in synthesizing musical rhythms with MIDI, that he might generate a more natural and satisfying performance.

Origin Of MIDI

MIDI was developed in the early 1980's. At that time, musicians were becoming increasingly frustrated at the incompatibility of synthesizers made by different manufacturers. Musicians wanting to gang together several synthesizers of different makes would likely have to rewire the various units to make them work together. Obtaining the necessary technical information was difficult because the various manufacturers lacked knowledge about their competitors' proprietary hardware designs--each knew how his

own system worked, but probably did not know enough to make it work with another system. (Boom, 1987 p. 11)

At the June 1981 National Association of Music Merchants (NAAM) show, however, the seeds of a universal musical instrument interface were planted. Men representing three major synthesizer manufacturers--Sequential Circuits, Roland Corporation, and Oberheim--discussed the possibility of a standard that would allow synthesizers of different manufacturers to communicate with each other. In the following months, Dave Smith of Sequential Circuits wrote a proposal for a Universal Synthesizer Interface (USI) to address the compatibility problem. (Boom, 1987 p. 11)

By the June 1982 NAAM show, the proposal had attracted the interest of more synthesizer manufacturers; these companies provided input regarding its specifications, and tacit agreement was reached on what was to become MIDI.

The first MIDI-equipped synthesizer on the market was Sequential's Prophet-600, first available in December of 1982. (Milano, ed. 1987) However, the *MIDI specification* was yet to be issued and compatibility was still a problem.¹

In August 1983, the MIDI 1.0 specification was settled at a meeting in Japan which included synthesizer manufacturers, Sequential, Roland, Yamaha, Korg, and Kawai. Production of MIDI equipment began in earnest thereafter. Another modification of MIDI came in 1985, with the issue of the MIDI 1.0 Detailed Specification. The detailed specification included a tighter definition of the uses of some of the continuous controllers and the requirement that manufacturers publish an explanation of their system-exclusive codes in their equipment manuals.²

Since the mid-80's, the MIDI acronym has become a familiar sight in both computer and media production trade publications. It has been promoted as a tool for professionals and as recreation for home computer and music hobbyists. Two trade organizations, the MIDI Manufacturer's Association (MMA) in America, and the Japanese MIDI Specifications Committee (JMISC) in Japan were established to ensure continued cooperation among manufacturers, so that

1. Cooper tells of a failed attempt to interface a Sequential synthesizer to a Yamaha at the June '83 NAAM show. Scoffers in the crowd dubbed the new interface "MUDI", for Musically Unusable Digital Interface. (Milano, ed. 1987)

2. The continuous controllers are used for real time modulation of the sound. Uses include, stereo pan, pitch bend, volume, vibrato, etc. System exclusive data can be used, among other things, to send the parameters of preset sounds from a synthesizer to a computer for on-screen manipulation.

the MIDI specification is maintained. These groups would also coordinate any discussions on revision of MIDI 1.0.

The last important development in the evolution of MIDI has been the Standard MIDI File format (SMF). It was adopted by the MMA in 1988. (Rubenking, 1991 p. 363) SMF allows musicians to exchange MIDI song sequences created from different software packages. It corresponds to the hardware compatibility already fostered by the MIDI specification.

What Is MIDI?

Simply put, MIDI is a data interface specially designed to communicate musical messages. An interface, in general, is a means of connecting two or more devices together; it defines how data is to be communicated within a particular system. For example, an interface may govern how fast data is to be transferred; the voltage level of the electrical signal; the means used to indicate the beginning and end of data types, etc. The MIDI interface may be looked at in terms of 1) electrical orientation; 2) communication protocol.

Regarding electrical orientation, the MIDI specification calls for a standard (5-pin DIN) connector. It also states the rate and mode of data transfer, 31.25 Kbaud, or, 31,250 bits of digital information per second, transmitted serially. Finally, it details how the electrical signal will be connected to the device.³

In terms of communication protocol the MIDI specification is music-performance-specific. In other words, its digital language is designed to describe musical events. Individual codes are used to identify the type and degree of various musical events. For instance, when a key is pressed on a MIDI keyboard, it will send out a Note-On message from its MIDI OUT port, followed by another message, called Velocity, indicating how hard (on a scale of 0-127) the key was depressed; finally, it will send out a Note-Off message when the key is released.⁴ Each of these messages is identified by its particular code.

A MIDI message is communicated in binary language. In a binary language, there are only two possible states: on or off, 1 or 0. A stream of MIDI data, then, is made up of

3. The specification calls for an optically isolated connection. This affords some protection from major damage to a device, should it be connected a defective device or connected improperly.

4. Some devices indicate Note-Off by sending out a Note On, Velocity=0 message instead.

thousands of binary digits, or bits, each of which has a value of either 1 or 0. These bits come one after another down a single path, because MIDI is a serial interface.

This data stream is divided into groups of eight bits, called bytes or words. A MIDI device will count off eight bits and recognize that as a byte, then count off eight more, call that a byte, and so on.

A MIDI message is typically made up of several bytes. The MIDI specification sets out two types of bytes: status and data. The status byte, coming first, will indicate the type of event (e.g., key depression, pitch bend, etc.) and the channel on which it occurred. This will be followed by one or more data bytes giving the degree to which the event occurred (e.g., which key was pressed and how hard, how far the pitch was bent). Whether a byte is a status or a data byte is indicated by the first bit of that byte. If it is set to 1, the device will interpret the following seven bits as belonging to a status byte; if it is set to 0, it will do the same for a data byte. This identification of a byte is made after the "start" and "stop" bits have been stripped off. These special, fixed status bits begin and end every byte. They are important to the business of serial data transmission, but do not carry any musical information. Thus, they are stripped away, and what is left is either a status or data byte, identified by the setting of its first bit.

Of particular interest is the ability of the status byte to denote the channel of the MIDI event. Usually, channels are conceived as separate pathways connecting two points. Recall that in MIDI there is only one electrical pathway, i.e., a single wire and its ground. The subdivision of this single path into 16 discreet channels is accomplished through the status byte: its last (or, more properly, lowest) four bits can represent a number, 0-15. This translates into MIDI channels 1-16.

This manner of setting the first bit of a byte to 0 or 1 in order to identify it has the effect of limiting the range of values that a byte can express. The succession of 1's and 0's that make up a byte can be expressed as an integer; for instance, 1 0 0 0 0 0 0 1 equals 129, and 0 0 0 0 0 0 1 equals 1. In the case of a status byte, where the first bit is set to 1, the range of values possible is 128 to 255; for the data byte, it is 0-127. Thus, the processor of a MIDI device can distinguish between them by a simple conditional comparison: if the byte's value is less than or equal to 127, it is a data byte; if it is greater than 127, it is a status byte. (Boom, 1987 p. 71) This coding scheme has implications for the expressive range of MIDI, since it allows a range of only 128 values that a parameter can vary in.

It is seen, then, that what is transmitted down a MIDI cable is not audio signal information; rather, it is digital information describing musical events. Only after MIDI data is fed in to a tone generator--a device that takes the MIDI event messages and converts them to sound--can the bits and bytes be heard as audible music.

A last note on the MIDI interface will help clarify the following discussion. MIDI is a *compound* interface; that is, it can handle both *event* and *timing* messages simultaneously. MIDI can be used to record and play back a performance on a keyboard, and it can also be used to synchronize playback of performances stored in several MIDI devices, such as a sequencer and a drum machine. Or, it can be used simply to link several synthesizers together, so that a note played on one will sound on all. Depending on the application, MIDI will function as either an event or synchronizing interface, or it may function as a combination of both.

Synchronization And Timing

In order to investigate the capacity of MIDI to render musical rhythm naturally, one must first understand how MIDI keeps time. A clock message is sent out in the MIDI data stream that synchronizes all devices in the system to a chosen master device. The MIDI Timing Clock message has relevance for such devices as sequencers and drum machines, when they must work in concert to reproduce a piece of MIDI music. In other situations, i.e., where musicians are using MIDI to link several synthesizers to produce a combined sound for live performance, the Timing Clock message is irrelevant and the devices ignore it.

The Timing Clock message is sent at a rate of 24 ticks per quarter note, as defined by the MIDI specification. The *tempo* of a piece of music is set by the user and is defined as the number of quarter notes per minute. Thus the rate of transmission in real-time for clock messages is dependent on the tempo of the piece being performed, and will vary in proportion to it.

The division of a quarter note into 24 parts was not an arbitrary choice. As with other aspects of the MIDI protocol, it has a musical foundation. A quarter note, or, "beat" articulated into 24 parts allows for an even division into shorter notes that corresponds exactly to the system of musical notation.

For instance, a quarter note divided in two becomes two eighth notes, each would have a duration of 12 Parts Per Quarter-Note (PPQ). Dealing with triplets, where a note value is divided into three equal parts, is possible: an

eighth note triplet--dividing a quarter note by three--has a value of $24/3$, or, 8. (see Figure 1)

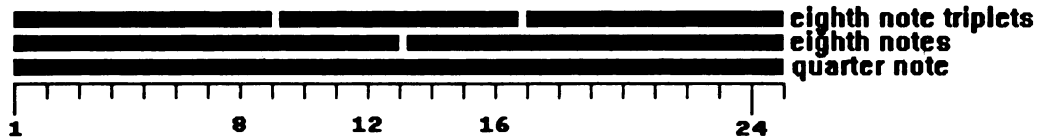


Figure 1. Twenty-four PPQ Quarter-Note Subdivided

Based on 24 PPQ, the smallest note MIDI can distinguish is a 64th note triplet, which has a duration of 1 PPQ. Since this and smaller values are fairly rare in music, the 24 PPQ designation is acceptable for synchronization purposes. But in representing human rhythm performances, a resolution of 24 PPQ is not fine enough. MIDI devices, however, are not necessarily limited to the 24 PPQ standard in representing musical rhythm. A closer look at the MIDI Timing Clock message will provide a useful understanding of the distinction between event and synchronization messages in MIDI. And this distinction will provide the basis for a clear view of the capacity of MIDI to represent rhythms naturally.

System Real Time Messages

The Timing Clock message belongs to a group of MIDI messages called *system real-time* messages. System real-time messages coordinate timing in the MIDI system. They are made up of a single status byte that does not carry a channel designation. System real-time messages are received on all channels, by all devices of the system capable of receiving MIDI data. System real-time status bytes may be sent within other messages to make sure they arrive at the intended time. For instance, a Timing Clock message could be inserted between the status and data bytes of a Note-on message. (see Figure 2)

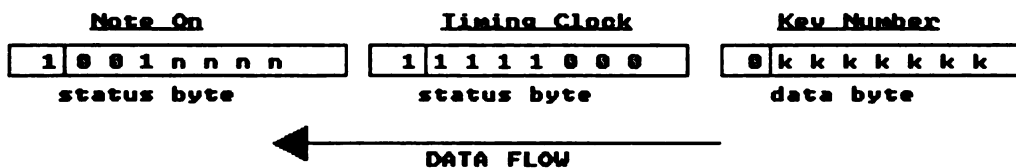


Figure 2. Timing Clock Message Inserted Between a Status and Data Byte

The master device will send out clock messages at 24 PPQ continuously, even while no other data is being sent. This allows the slave devices to calculate the tempo of the piece the master device is about to play. The calculation is made by measuring the real-time rate of clock messages in seconds. If Timing Clock messages are arriving at 24 per second, that equals one quarter note per second, or a tempo of 60. When the system real-time message, *Start*, is received the slave device will start playing its sequence in perfect synchronization with the sequence playing on the master. Any other messages emanating from the master's MIDI OUT port will be ignored by the slave: the MIDI link, in this case, is functioning only to synchronize the tempos, start and stop times of the sequencers in the various MIDI devices.

This sort of a master/slave hook-up is rather typical. It may include a personal computer (PC) as the master sequencer and a keyboard and drum machine as the slaves. The keyboard would act on note information coming from the PC, using it to control its tone generators. The drum machine would act on the PC's synchronization messages, starting its own sequencer at the appropriate time and running at the correct tempo; the drum machine's *internal sequencer* sends note information to the tone generators inside the drum machine. The audio outputs of the keyboard and drum machine could be combined using an audio mixer to create the total sound.

It is important to observe that the quality of the drum rhythm produced with this set-up is dependent on the *beat resolution* of the drum machine, and is not dependent on the 24 PPQ standard or the sequencer. Of course, a high-quality drum machine with a high PPQ is possible. But generally drum machines have PPQ values less than 100, and synchronization requires conversion where the drum machine operates at a PPQ other than 24.⁵ More critically, they are likely to have comparatively limited memory space and difficult editing features.

The hook-up described above demonstrates the utility of the 24 PPQ standard in synchronization. Furthermore, the sequencers can stop or start at any of the 24 parts of the quarter note, which correspond to musical notation. Therefore, the synchronization is musically coherent.

However, with this arrangement, coordinating changes in the drum part with changes in instrument parts involves

5. Drum machines typically operate at 24, 48, or 96 PPQ internally. Synchronization problems can result when the beat resolution of the machine does not match that of the synchronizing signal.

editing of two different sequences; and the PC's sequencer is likely to be capable of much longer sequences without repetition than the drum machine's. This mismatch may limit the ability of the drum machine to follow musical events in the instrument parts, because the drum pattern would have to repeat at some point to match the length of the longer sequence.

There is another type of hook-up, in which synchronization is not necessary, that can prove more musically versatile. Instead of utilizing the internal sequencer of the drum machine, it is possible to command a tone generator containing both drum and instrument sounds, or an instrument tone generator and a drum tone generator separately, from a single sequencer. The musician would write a single sequence for the whole performance, including drum and instrument parts, on a PC. Then, the software sequencer in the PC would send out note information via MIDI to the tone generator(s) and the performance could be realized, without the necessity of synchronizing several sequencers.

This sort of arrangement provides the greatest convenience in composing and editing a piece of MIDI music because the whole sequence can be accessed from one point. Moreover, it allows the PC to set *the beat resolution*. This gives the ability to utilize PPQ's much higher than 24, 48 or 96, thus increasing MIDI's capacity to represent rhythms more naturally. To understand how this works, one must know how a MIDI sequencer records and plays back music.

Recording and Playing Back on a MIDI Sequencer

In order to record a MIDI sequence with a sequencer, the event data must be referenced in time. In other words, it is not enough to remember *what* occurred, the sequencer must also remember *when* a given event occurred. To accomplish this, the sequencer employs an *event clock* to make a list of the type of incoming event (e.g., Note-On, Pitch Bend, etc.) and the time that it occurred.

Here is how it works: The sequencer's microprocessor has access to a counter that it uses to count the ticks of the event clock.

"When a message is received, the processor stores the message along with the counter number. Then it resets the counter to 0 and waits for the next message. The counter starts counting ticks again. When the next message arrives, the message and the new counter number are added to the list. The counter is reset and the process repeats. The resulting list contains not only the list of performance

messages in the order that they occurred, **but the number of clock ticks between each message.**" (De Furia, 1986 p. 20)

This is how a sequencer records a MIDI performance. This list of events can then be written in a sequential file and saved on a computer disk. A sequential file is read in a first-in, first-out fashion. This makes playback fairly straight-forward:

"When the performance is recalled, the processor reads through the list, one message at a time. Using the Event Clock as a timing reference, it waits for the number of ticks stored with a message and then transfers the message to the appropriate voice. It then waits for the number of ticks stored with the next message and transfers it...Slowing down or speeding up the clock speed will alter the tempo of the music as it is played back." (De Furia, 1986 p. 21)

At the user level, a MIDI sequencer appears much like a conventional tape recorder for recording and playing back pieces of music performed on a MIDI controller, such as a keyboard. Regarding the representation of rhythm, however, there is a critical difference between a "MIDI recorder" and an analog tape recorder. That is, time is fluid on an analog tape: the audio signal can vary in time continuously, with faithful representation. Contrast this to MIDI, where time is marked at intervals. Events falling between those intervals cannot be recorded as they occurred in the original rhythm context. De Furia lays this problem out plainly:

"This method of [MIDI recording] is accurate to the **nearest clock tick**. When an event occurs in between clock cycles, the counter number will be for either the click before or the click after the event. When the performance is replayed, that event will be either a little early or a little late." (De Furia, 1986 p. 21)

This is one of the prices we pay for "digitizing" music reproduction with MIDI. The timing of MIDI events can never be fluid, as with audio tape. Thus, performance of a sequence of notes by a MIDI sequencer or drum machine must be viewed as fundamentally mechanical. The mechanical nature of sequenced MIDI performances would not be a problem, however, if human performances did not deviate from prescribed note values. After all, the 24 PPQ division corresponds in exact proportion to the traditional system of music notation--down to a 64th note triplet. However, research has shown that musicians rarely match prescribed note values for duration.

A way to reduce this mechanistic quality is to increase the beat resolution of the sequencer. This can be done by

increasing the rate of the event clock relative to the tempo of the piece. If tempo=60 (one quarter note per second), increasing the clock rate from 24 to 240 Hz would provide a one-hundred times gain in beat resolution. The event messages are then recorded, edited, and played back in reference to this faster event clock. *Timing Clock* messages can still be sent out at 24 PPQ, by sending them out once every 10 ticks, instead of on every tick. Therefore, a sequencer's beat resolution may be much higher than the 24 PPQ synchronization standard.

It is probable that with high enough PPQ values, a phenomenon similar to persistence of vision will make MIDI rhythms seem fluid enough to satisfy an audience.⁶ This is based on the empirical threshold of human perception for differences in the time spacing between consecutive notes. Seashore (1938, p. 91) reports this to be 10 msec for a "very fine musical ear," and as high as 100 to 200 msec for ears with a duller sense of time discrimination.

At 240 PPQ, the duration of 1 PPQ is equal to 2.08 msec, at a tempo of 120; at tempo=90, 1 PPQ=2.7 msec; for tempo=60, 1 PPQ=4.1 msec. These tempos range from moderate to moderately slow. For each, the duration of 1 PPQ--the smallest time value a sequencer can discriminate--is well within the 10 msec threshold for a "fine musical ear." Of course, for tempos higher than 120, the 1 PPQ value would be even less. The beat resolution of the sequencer used in this investigation is 240 PPQ.

MIDI Delay

Because of the way the MIDI protocol is set up, each MIDI channel can carry a lot of information. Theoretically, a single channel could have 128 notes sounding all at once. These notes could also be modulated with any or all of the 64 continuous controllers and pitch bend. Theoretically, all 16 channels could transmit all of this at the same time. This sounds very impressive, but in practice there are limits on MIDI's data handling capacity that fall short of these theoretical limits.

Since MIDI is a serial interface, data moves through the cable as a stream of individual bits. Furthermore, it takes ten bits to make up a single byte and multiple bytes to comprise many MIDI messages. Therefore, there is a time differential from the moment a message is sent from a MIDI OUT port to when it is recognized by the receiving device as a message and then acted on. For a single byte the interval

6. Persistence of vision is what makes motion pictures possible. They are made up of discreet frames, yet are perceived as fluid motion.

is quite small: 320 microseconds. However, pressing and releasing one key on a keyboard will generate up to six bytes, three for the Note-On message and three for Note-Off. In this case, it takes MIDI 1.92 milliseconds to communicate that a single key has been pressed and released.

MIDI's serial design has built into it a certain lack of responsiveness. In polyphonic music, there are likely to be two or more tones sounding at any given moment. Also, multiple tones will often be struck at the same instant. When polyphonic music is conveyed via MIDI, it is serialized. In other words, no two tones can ever be produced with perfect synchronization--they all come in a line. The notes of a MIDI chord, then, though they may seem to sound simultaneously, are in fact sounding in rapid succession.

To make matters worse, it takes the processor inside a MIDI device a certain amount of time to process the MIDI data, leading to yet another delay. For example, when a key is pressed on a MIDI keyboard, it takes between 5 and 7 milliseconds until the Note-On message is present at the MIDI OUT port. If a musician routes that MIDI out signal to the MIDI IN port of another synthesizer, he will add another 5 to 7 millisecond delay to the time between when he presses the key on the first synthesizer to when he hears it sound on the second. This added delay is how long it takes the processor in the second instrument to transfer the message to its voices. (De Furia, 1986 p. 52)

While this sounds discouraging, MIDI's high baud rate often makes interface delays negligible. At other times unacceptable delays are due mainly to the processors within the various MIDI devices. Ten notes of polyphony can be sent through the interface in 6.7 milliseconds. (De Furia, 1986 p. 54) Unless another event is to follow in less than 6.7 milliseconds, the ten notes have enough time to get through before the next event.

To arrive at the 6.7 msec figure, one begins by dividing 31250 by 10. This yields 3125, or the number of bytes MIDI can transmit in one second. Dividing this by 1000 gives the number of bytes per millisecond, or 3.125. Assuming it takes three bytes to sound each of the ten notes in the chord, this means that 30 bytes must be divided by 3.125 to get the number of milliseconds it will take to transmit the chord. The figure obtained is 9.6 msec.

The reason for the discrepancy between 9.6 and 6.7 msec lies in something called *Running Status*. Most MIDI devices will implement Running Status to reduce the number of status bytes in the MIDI data stream. For a group of notes, such as our ten note chord, the receiving device will read the first Note-On message and go into Note-On mode. If the

transmitting device implements Running Status, it will send out a Note-On message (with appropriate channel designation) followed by two data bytes: one for the key number, one for the velocity. Then, instead of sending a Note-Off message to terminate the note's duration, it will send a data byte for that note's key number, followed by a Velocity=0 data byte. If the receiving device receives no new status byte, it assumes that the following data bytes apply to Note-On status, and it remains in Note-On mode. For the remaining nine notes in the above example, only the key number and velocity data bytes would be sent. Provided that no other type of status byte is sent, an infinite number of notes can run under the status of a single Note-On status byte.

In the case of groupings of notes or consecutive notes the Running Status feature will reduce the data flow by about one-third, by eliminating Note-On status bytes from all but the first note. To adjust the above 9.6 msec figure for the effects of Running Status, first multiply by one-third (.33). This yields a product of 3.17. Subtracting 3.17 from 9.6 will in effect delete all status bytes from the 9.6 msec time frame. The result is 6.43. Adding .32 to this figure re-introduces the first Note-On status byte to the elapsed time, yielding the 6.75 msec figure.

It is important to observe that 6.7 msec is the time it takes MIDI to send ten Note-On messages. It requires additional data to turn those notes off. If one wanted a rapid succession of ten note chords, then, it takes about twice as long (13.08 msec) for MIDI to transmit the information for each chord. The 13.08 msec figure is arrived at by doubling 6.7 to add in the key number and velocity = 0 data bytes for turning the notes off, and then subtracting .32 to remove the duplicate Note-On status byte.

Even with MIDI's high baud rate and the advantage of Running Status, unacceptable delays can crop up. If there are a large number of notes to be played simultaneously in a crowd of fast notes, delays may spread the simultaneous notes out perceptibly. A lot of real-time modulation from the continuous controllers would exacerbate the situation. Finally, faster tempos increase the volume of data per unit of time, possibly introducing delays.

The implication of MIDI delay for the representation of rhythm is that it may distort the timing of MIDI rhythms. In fact, MIDI rhythms are always distorted. Take a situation where a bass drum hit, a four note organ chord, and a trumpet note are all supposed to strike on the first beat of a measure. MIDI will spread out their onsets in time. The six tones can never *sound* simultaneously. While it is doubtful that human performers ever play in perfect synchrony, it is clear that at times *they may*. Such is never an option in MIDI synthesized performances.

Therefore, at whatever minute level, MIDI is introducing a certain amount of systematic variance in rhythm due to the serial interface design. In a sequenced MIDI performance, then, where drum sounds and instrument sounds are travelling on the same MIDI cable, will the drum rhythm remain coherent within itself? Or, is it subject to being buffeted about as instrument sounds compete to occur at the same instant as drum sounds?

MIDI sequencers represent rhythmic elements in their sequences as though they are going to occur at precise intervals. In performance, however, the rhythms vary from their representation. This is akin to a human performer deviating from the prescriptions of music notation. If the goal is to produce rhythms with MIDI that are more human sounding, then it must be ascertained 1) how and to what degree sequenced MIDI rhythms vary from their representation; 2) how and to what degree human performers vary from the prescriptions of music notation.

The amount of precision required in measuring the difference between a sequencer's representation and its performance of a rhythm pattern is tempered by the 10 msec threshold of human perception of rhythmic differences. The maximum delay can be calculated based on the number of polyphonic notes at any moment in the sequence. For instance, 15 notes can be transmitted in 9.968 msec. If a drum note is to be struck simultaneously with fourteen sounds or less, its maximum delay will fall within the 10 msec threshold. This means that if the MIDI musician keeps certain limits in mind, he may ignore MIDI delay as an unwanted source of systematic variance in rhythm synthesis.

A final observation is necessary on the processing lag mentioned above. While the MIDI delay can be calculated based on the 31.25 Kbaud standard, processing lag will vary from device to device, based on the quality of design. Most of the delay in a MIDI system is due to processing and not the MIDI interface. (De Furia, 1986 p. 53) In a system where the drum and instrument sounds are resident in a single tone generator, the drum and instrument sounds will experience the same processing delay. This will appear as a slight lag between when start is pressed on the sequencer and the first sound is heard. While this delays start time, it does not affect the internal rhythms of the piece.

If the drum sounds and instrument sounds are in different tone generators, however, the internal rhythms of the piece may be distorted by the differential processing lag of the various devices. Nevertheless, if the difference is known, many sequencers allow individual tracks to be shifted slightly in time. The differential lag could be compensated by shifting the rhythm track so that it is

brought into *aural* synchronization with the instrument tracks.

MIDI AND RHYTHM RESEARCH

Researchers have been interested in human rhythm response since the latter half of the nineteenth century. Since that time, improvements in technology have aided in the registration, analysis and synthesis of human rhythm performance. Better equipment is always desirable because it provides more accurate measurement, faster and more complete analysis, and tighter scientific control over music parameters. It is part of this investigation to see how useful MIDI is in registering performances. Another objective is to see how well MIDI can synthesize a human rhythm performance. The study will also explore the convenience of using the Standard MIDI File in computer analysis of music performance.

Since MIDI is seeing increasing use in media production as a substitute for human performers, methods discovered by research that make MIDI performances sound more human will have commercial applicability. A better sounding MIDI sequence will make a better contribution to whatever production it is used in. With a clear picture of what MIDI can and cannot do, media practitioners can get the most out of this new technology.

Also, MIDI is offering unprecedented power of expression to independent composers, song writers, and musicians. The ability to compose and record with a full palette of digitally sampled sounds within one's own home was only a dream ten or fifteen years ago. Today, this capability is within the reach of many, and the price of personal computers--one of the most costly components in a "MIDI studio"--continues to decline as their power increases.

For the musician operating on a small budget, it is essential to get the best sound out of every equipment dollar. A big part of this sound quality maximization is understanding how the equipment works and what tricks, if any, will make it work better. For instance, what good is it to have 16 bit, digitally sampled sounds in a MIDI studio when the sequencer is no more rhythmically sophisticated than the music box on the mantle? In a professional sense, not much. Research into MIDI's rhythmic capacity could benefit small budget musicians, composers and song-writers, provided it is applicable to equipment typically used by them.

The basic goal of this investigation is to establish a clear view of MIDI's rhythm performance capacity, and to see if any improvements can be made. Specifically, it wishes to discover 1) how well MIDI can capture natural human rhythms; 2) what systematic variance is used by a drummer in the

performance of certain rock and roll rhythms; 3) can the observed systematic variance be applied to a purely synthesized rhythm.

In addition to understanding the technical facts, one needs to be apprised of the existing body of literature on human rhythm performance and perception in order to conduct a responsible investigation of the topic. This is helpful in establishing measurement methods and interpreting results. The studies hint at what the registrations are likely to reveal. Also, they record idiosyncrasies in human perception of rhythmic elements that, if unheeded, might confuse interpretation of the data.

Musicology Research on Rhythm

All of the studies begin with the traditional system of music notation as a framework, and compare to it empirical observations of performance. Therefore, any variance found in these studies is conceived as variance from the so-called "rational-mechanical" norm, i.e., the notational system.⁷

The studies have several important facets. First, the method of registration is noteworthy, because it is the chief limiting factor in data collection and analysis. This hinges mostly on the degree of precision attainable in measuring the times of rhythm elements. This level of precision establishes a reference to which MIDI registration can be compared. If, for instance, good results have been achieved using another registration method with an accuracy of +/- 10 msec, this provides a window of opportunity for MIDI registration.

Second, later studies provide evidence of systematic variance in rhythm performance. This is enough to condemn MIDI rhythm performances structured around the rational-mechanical norm. Nonetheless, if MIDI is capable of expressing the level of variance found, evidence of systematic variance may be put to use in modifying MIDI performances to make them more human sounding.

Third, some studies suggest what it is that makes a human being perceive a series of sounds as a rhythm. The concept of accent is important, as is *grouping*. Systematic variance in several note parameters is used to indicate accent and grouping. When trying to synthesize drum rhythms, *the synthesis of such variance is essential to a human feel.*

7. The term "rational-mechanical norm" was coined by Bengtsson and Gabrielsson (1980, p. 257).

Finally, studies indicate that human perception of rhythm is somewhat anomalous. For example, increasing the amplitude of a note may make its duration appear to increase, as well. This suggests that drummers may use *effects of appearance* caused by the interaction of different note parameters on human perception. Consequently, the registration must capture all the relevant parameters, and the analysis must recognize the possible use of such effects.

Specific Studies

Sears in 1902 conducted what was perhaps the first study into musical rhythm to use electrical technology for registering a performance. He registered organ performances of hymns by means of an organ with wired keys and a kymograph. The kymograph was simply a rotating drum with paper on it, over which were poised several pens each connected to an electromechanical apparatus that would move it when its respective keys were depressed. Several keys were wired to each pen. In addition, a clock reference was sent to one pen and also written on the paper.

Analysis of the kymograph record was a laborious task because each pen line was a record for several keys. Moreover, each performance had to be short to avoid overwriting the record paper on the drum. Nevertheless, Sears discovered definite variance from the rational-mechanical norm. He analyzed the performances for tempo variance in the overall performance and down to the measure level, variation in the length of measures, the duration of notes, way of accenting, etc.

Regarding accents, Sears found a tendency to lengthen the duration of the accented note:

"It is evident from the foregoing that accented notes are often longer than unaccented notes of the same denomination, but it is also evident that this tendency is not present in all cases and with all players." (Sears, 1902 p. 46)

On some hymns there was a marked tendency to lengthen accented notes; on others, the tendency was weaker. Also, there were differences in direction and degree between various players. This suggests that the method of accenting is *context-specific*, because the device of lengthening the accented note was used more or less depending on the hymn played. Also, accenting may be *style-specific*, in that some players show the opposite tendency, that is, they shorten accented notes. Indeed, this may be one aspect of personal style.

Sears quotes an interesting passage from an 1895 French study by Binet and Coutier:

"In relation to the accentuation of single notes these researchers found that a tendency exists-- 1. to separate the accented note from the preceding note, 2. to tie or slur the accented note to the following note, 3. to increase the length of the note accented as if this were equivalent to an increase in intensity, 4. to increase, especially in rapid playing, the intensity of the notes which follow the note accented."

Sears' work provides a historical backdrop for one aspect of the current investigation, e.g., analysis of performance registration. Furthermore, it gives some glimpses into the ways in which performers accent tones. However, because he used an organ for his registrations, the role of amplitude (velocity, in MIDI parlance) in accenting was something Sears could not investigate.⁸ A study that looked at the interplay of amplitude and duration in rhythm was Woodrow's in 1908.

Woodrow did not analyze performances; rather, he generated sequences of tones, manipulated the intensity, spacing and duration of certain tones, and tested the effects on listeners. He used a primitive, motor-driven rotary switch-like apparatus, which was carefully monitored for accurate calibration. His basic question was, what is it that makes the listener perceive a series of tones as a rhythm and not merely a series of tones?

Woodrow started his experiments by trying to find the "indifference point" in a series of tones. This was the point where the listener did not perceive any rhythmic grouping in the series--where it sounded void of rhythm. From the indifference point, systematic variance was introduced in the intensity, duration and spacing of tones to produce a sensation of rhythm. Also, different types of rhythms can be produced by manipulating these parameters:

For spacing,

"It is possible to pass from one rhythmical grouping to another by changing the relative duration of the intervals between the sounds. Thus, a trochaic rhythm, that is, one that is composed of groups of two sounds each, the louder sound beginning the group, may be changed to an iambic rhythm, one in which the louder sound ends the group, by increasing the interval immediately following the louder sound or by decreasing the interval immediately preceding it."

8. Traditionally, an organ cannot vary the amplitude of individual notes according to how hard they are struck.

For intensity,

"When the intervals are equal and every second stimulus the stronger, the rhythm is trochaic, and when every third is the stronger, dactylic. That is, a regularly recurring difference in intensity exerts a tendency towards rhythmical groups with the more intense sound at the beginning."

For duration,

"With an increase in the ratio of the duration of the longer sound to that of the shorter, there is an increase in the tendency of the longer to end the group or a decrease in its tendency to begin the group." (Woodrow, 1908 pp. 63-64)

Woodrow's results shed light on the "grouping" concept. Drummers probably manipulate spacing, duration, and intensity to group the sounds of their drums into musical measures and phrases, and to emphasize structural aspects of the particular piece. Therefore Woodrow's findings suggest things to look for in analyzing drum registrations. However, one must apply his results advisedly, because they were obtained from non-musical performances.⁹

In the late 1930's, Carl E. Seashore and his colleagues at the University of Iowa did some extensive investigations into music performance. They developed something called the Iowa Piano Camera for registering piano performances. The camera took pictures of the key action on a wide roll of film. Its accuracy was reported to be 10 msec; it was capable of recording how hard a key was pressed, and when pedaling was used.

Seashore used professional pianists playing classical pieces to get his data. He viewed deviation from the rational-mechanical norm as essential to quality music:

"It is often stated that great accuracy in the hearing and the performance of rhythm is not of much consequence because there is such great irregularity and license in the rhythm of even the best music. This notion is based on the assumption that rhythm should occur in metronomic time. The musician, however, knows that his artistry lies not in maintaining a rhythmic pattern in even time, but rather in the hearing and making of artistic deviations in the pattern. This is a far more strenuous demand than a demand for the setting of the pattern in even time. It is the delicate varying of pattern interpretations that puts life into music." (Seashore, 1938 p. 137)

9. For instance, creating a trochaic rhythm by increasing the intensity of every second tone may suffice to achieve grouping, but it would probably have a monotonous musical effect.

In other words, deviation from prescribed note values is no accident, but requires an acute sense of rhythm. Seashore's pianists "put life" into the scores in various ways. Some interesting findings regarded accents. Note intensity (how hard a note is hit) was not found to be essential to accent; rather, altered duration and delayed entrance of the accented note were the consistently used devices.

Vernon, another of the University of Iowa group, studied performances captured on piano rolls. He was looking at chord asynchrony, or, the degree to which a musician serializes a chord to emphasize its structural aspects or one of its tones. Vernon's work is noteworthy in that he formulated performance "rules" for chord asynchronization. Further, he claimed to have confirmed some of these with statistical analysis.

By far the most systematic and detailed studies of rhythmic performance have been conducted in Sweden, at the University of Uppsala. Begun by Ingmar Bengtsson in the late 1950's and carried on by Alf Gabrielsson, this ongoing investigation into musical rhythm has resulted in many published studies.

The Uppsala studies branch out in two areas: analysis of performance, and listener perception of musical rhythm. On analysis of performance, a long series of papers have come out. The main problem of these investigations is:

"How do the musicians actually play to bring about the intended/desired rhythm characteristics?" (Gabrielsson, 1979 p. 83)

Their basic hypothesis is that "good/typical performances of music associated with specific rhythm characteristics," such as swinging jazz, "are characterized by certain systematic variations in relation to a 'rational-mechanical' norm for the performance." (Gabrielsson, 1979 p. 84)

As a means to examining this systematic variance, a device was designed which would graphically record the wave forms of audio recordings of sound sequences. It provided registrations with an accuracy of 10 msec or better. In various experiments, different performers were asked to play the same short melodies or rhythm lines on such instruments as the piano, flute or bongo drum. The registrations of these performances were transferred to computer for analysis. Ultimately, a data set would be put together that contained, for each performer, the systematic variance applied in his or her rendition of a particular piece of music. The researchers would then try to correlate the numerous renditions using factor analysis to identify a few

fundamental "performance types." Usually, this factored out two or three typical ways to perform a piece in terms of systematic variance from the notation values on the score.

The Uppsala studies are a great contribution to the method of capturing and analyzing systematic variance in rhythm performance. They have come up with some more detailed results of performance behavior:

"A sequence of two eighth-notes within a beat is seldom performed with equal durations for each of them. Short-long relations (S-L) appeared generally for pianist A and predominantly for pianist B. For the percussionist there were more varying results."

"A sequence of an eighth note followed by two sixteenth-notes was performed with long-short (L-S) relations on the eighth-note level but with S-L relations among the sixteenth-notes by the percussionist and often by pianist B."

"There were striking deviations from notation norms in connection with syncopations (the percussionist). One such phenomenon was a relative prolongation of the eighth-note values and a relative shortening of the sixteenth-notes at the beats where the syncopations occurred."¹⁰

"At the performances of [rhythmic repetitions of a single tone] the highest peak amplitude invariably occurred for the first sound event of the measure and it seems clear that this is intentional for the sake of a perceived accent on that position." (Gabrielsson, 1974 p. 72)

In another study, the Uppsala group was able to document the relative shortening and lengthening of beats within a measure. This systematic variation was said to be essential to achieve the proper rhythmic feel of a Viennese waltz. (see Bengtsson and Gabrielsson, 1975)

While these results have limited applicability to rhythm synthesis, they support some of the findings of earlier researchers. Also, they help direct analysis when used as starting points in looking for systematic variance. If anything can be found wanting in these studies, it is *generalizable results*. Perhaps the researchers are carefully approaching the point where they can make some generalizations. Whatever the case, a body of *performance rules/guidelines* must be generated, if such studies are to have significant value for the music synthesist.

10. "Syncopation is the displacement of either the beat or the normal accent of a piece of music."--The Oxford Companion to Music. 10th Edition: p. 1002.

Another group of researchers at the Royal Institute of Technology in Sweden have put together a system of performance rules that can be used in music synthesis. These rules depend on the musical context within a given piece of music. A particular rule is applied if certain conditions are met. The rules control such parameters as duration, amplitude, vibrato amplitude, relative frequency deviation. They are expressed as equations, which allows accurate application.

A recent study (Friberg et al, 1991) describes an audience test of the rule system. The panel was made up of professional musicians and composers. They listened to alternate versions of computer synthesized piano music, some versions applying the rules, others presented "deadpan" (no deviation from the rational-mechanical norm). A computer program was used to automatically modify the music sequence according to the rule system. The results were very strong in favor of the rules-modified versions. (Friberg et al, 1991 p. 53)

IMPROVING SEQUENCED RHYTHM QUALITY

It is clear at this point that a human rhythm performance is temporally complex. It is not static, but dynamic in relation to the rational-mechanical norm. It is little wonder, then, if machine-produced rhythms are easily recognizable as such, and that listeners so often find these synthesized rhythms unsatisfying. They simply do not match expectation.

This problem of mechanical-sounding synthesized rhythms has been addressed in several ways by software authors, equipment manufacturers, and computer music practitioners. These attempts have had varying results. It is a challenging problem because the method of correcting the mechanical feel must be general enough to work on many possible musical genres, yet specific enough to render each genre its characteristic style. The value placed on natural sounding sequenced rhythms is reflected in the complexity of a given solution.

The emergence of a "humanize" function in some sequencer software is an acknowledgment of the mechanical feel problem. The humanize function will introduce variance to certain note parameters to approximate a human performance. While this seems an excellent idea, a world of uncertainty opens up when trying to decide what variance will be introduced where. This reservation is dealt with by applying random variance to selected note parameters over a chosen span of notes. For instance, in the Master Tracks Pro (TM) sequencer program used for this investigation the operator can highlight a range notes or measures, or even the whole piece, for "humanization." Next, the operator decides what note parameters will receive the random variance. Available parameters are note start time, duration and velocity; any or all may receive the variance. Last, the operator determines how much variance each parameter will get by specifying the range within which it can vary. A press of the "enter" key and it's all done.

Such a humanize function has limited utility in that it is fast and somewhat controllable. Used with other methods it can be even more effective. The short-coming, however, is in the random variance. Musicology research has shown that performers vary from notated values in *systematic* ways. There will be random variance in a human performance, too. But this is the result of how refined the musician is technically; whether or not he happens to make any mistakes in playing; and, if there are any instrument design limitations that cause playing inconsistencies. (Bengtsson & Gabrielsson, 1980 p. 257) Therefore, the humanize function, inasmuch as it introduces only *random* variance, causes the performance to vary in the least functional way. However,

if a humanize function could be expanded to include systematic variance, it would be a convenient way to reduce the mechanical feel of MIDI sequences.

Some drum machines employ a so-called "swing function" to reduce the mechanical quality of their rhythms. A line of drum machines by the Roland corporation that include a swing function are their *Human Rhythm Composers*. The swing function is based on the idea that drummers will systematically delay certain notes within a given drum pattern to create the swinging style of rhythm. The Roland machines allow the user to select one of several "Swing Point" settings for a rhythm. The swing points are those notes in the repetitive rhythm pattern that will always be delayed. For instance, one Swing Point setting delays the second and fourth quarter notes of a 4/4 measure. (see Figure 3)

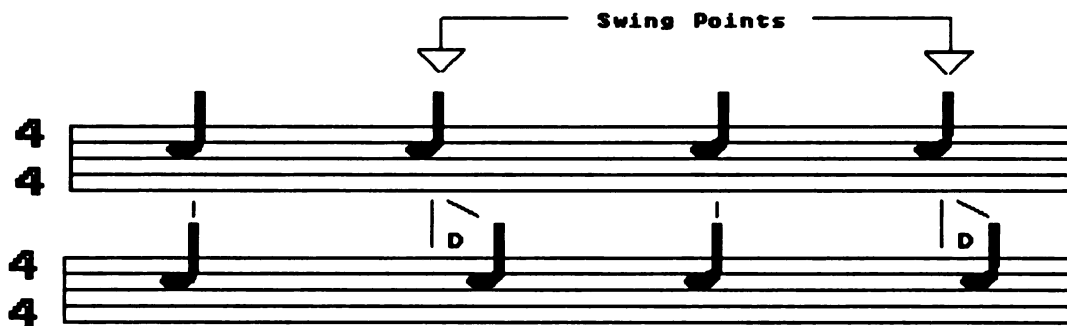


Figure 3. Sample Swing Point Setting for the Roland R-8 Human Rhythm Composer

The user may also set the amount of delay applied at the swing points. On the Roland machine this may vary on up to twenty-three levels, depending on the Swing Point setting. A model R-8 machine was obtained to assess the operation of the swing function. By recording its patterns via a MIDI link into the PC sequencer the amount of delay could be measured. It was found to change in increments of 10 PPQ at 240 PPQ resolution. The actual delay time, of course, depends on the tempo at which the pattern is being played. Once the delay and Swing Point are set, the machine will apply this systematic variance to the pattern every time it is repeated.

The swing function is an improvement over the humanize function, because it applies systematic variance instead of

random. However, a swing function such as this will swing mechanically, inasmuch as it applies the same amount of systematic variance at the same points on each repetition of the pattern. It is useful because it allows the drum machine to play patterns that are based on temporal relationships between tones that the notational system is not capable of representing. But a swing function alone is not sufficient to cover all the possibilities for systematic variance in a drum pattern. Thus, while it is closer to the goal of approximating human rhythm performance, it still lacks the necessary flexibility to accomplish the task.

Another possible way of making synthesized rhythms more human-sounding is a real-time performance interface that allows a performer to modulate the synthesized rhythms in time, imposing a human feel on a pre-composed structure. A device called the *Boie Radio Drum*, developed by Bob Boie at Bell Laboratories, does just this. The Radio Drum works in conjunction with the *Conductor* program created by Max Mathews. Together, they facilitate real time control of many facets of a computer-generated performance, including micro tempo. (Boulanger, 1990 pp. 34-39)

The Radio Drum is made up of two mallets with tiny radio transmitters in the heads and a matrix of receiving antennas. The receiving antennas are situated so as to set up a radio plane in their midst. This plane constitutes the Radio Drum's head. When the mallets are moved through this invisible plane, the position and velocity of the imaginary strikes are computed based on the signal strength of the respective transmitting mallet heads as received by the variously positioned antennas. The Drum reacts to continuously to variations in the x, y and z axes.

This information is interpolated and fed into the computer where it creates near instantaneous changes in chosen musical parameters, according to the motions of the performer's mallets. The parameters under the performer's control are such as dynamics, tempo, timbre modulation and accenting. What parameter is varied depends on where on the imaginary drum head the performer strikes.

The *Conductor* program is a custom sequencer that plays back a composition while reacting to the output of the Radio Drum. It can also record the gestures of the Radio Drum's performer. Thus, a performance, consisting of the note information in the sequencer and the recorded manipulations of the performer, can be stored, edited and reproduced.

The Radio Drum is the most dynamic solution to the mechanical feel problem in synthesized rhythms. Unfortunately, it may also be costly and is not commercially available. A tremendous amount of computer power is required to create the Drum, and to interface it to the

sequence. Such a system as yet belongs to the academic and avant garde computer music set. Furthermore, it would not necessarily employ the MIDI interface.

The concept of the Radio Drum, however, could probably be utilized on a lower level and at greatly reduced cost. The concept of real-time modification of computer-generated rhythms was in practice as early as 1971, as an adjunct to Leland Smith's *SCORE* program. The *SCORE* program used two telegraph keys to register the real-time input of rhythmic modulation from a performer. This information would then alter the existing rhythm in the piece for the next playback. (Smith, 1972 pp. 7-14)

A readily available way to introduce systematic variance to a sequence, which is highly accurate and flexible, is *individual note editing*. Most, if not all, sequencer software packages allow the user to access and edit the values of the various parameters that apply to individual notes. For drum rhythms, the relevant parameters are start time, duration and velocity.

Individual note editing is accurate down to the resolution of the sequencer. That is, values are edited in the smallest increments the sequencer can distinguish and reproduce. This gives the musician full control over the sound of the sequence. The trade-off is that this method can be very slow, especially without guidance on what notes should be edited and to what degree in order to produce the desired rhythmic effect. Nevertheless, it is one of the most cost-effective, accurate remedies to the mechanical feel problem, and it is probably available to every MIDI musician using a PC-based sequencer.

If a musician's sequencer allows individual note editing, it should be possible to reduce the mechanical quality of his sequences by altering the start times, duration and velocities of certain notes. When that sequencer has a high beat resolution, the chances for a satisfactory result are even better. The significant problem is knowing which notes to alter, in what direction, and to what degree.

If each piece of music were *absolutely unique*, knowing how to alter it would be impossible and the process would be reduced to trial and error. Fortunately, pieces of music, especially those within the same genre, will have some of the same characteristics. Taxonomically, it is possible to break music down into several constituent styles, analyzing in terms of melody, rhythm, harmony, counterpoint, instrumentation and orchestration, and form. Thus, people talk of classical, jazz or rock and roll music. Below this is another level, the "style-species," where a genre is divided into subcategories. Here is where bebop and swing,

for instance, are distinguished from one another within the musical genre of jazz. (Haydon, 1941)

If a piece of music is to belong to a specific style, this implies it must conform to the structural forms that define the particular style. Its structure will be in some way similar to other pieces written in that style. Rhythm is one of the factors that defines a given style. Therefore, common rhythm characteristics will probably be found between many pieces falling under a certain style. Inasmuch as systematic variance contributes to style-defining rhythms, it is reasonable to assume that, where there are structural similarities between two pieces of music, there will also be similarities in any systematic variance applied to those structures.

A simple example will illustrate this. The feel of a drum pattern for high-hat (H-H), snare (S) and bass drum (B) can be changed by altering the duration ratios of the eighth-note figures for the high-hat. The ratios will be altered relative to one beat, or, the duration of a quarter note. We begin where the ratio per beat is 1:1. Each eighth note occupies 50% of the beat. This produces a balanced, even-sounding rhythm (see Figure 4).

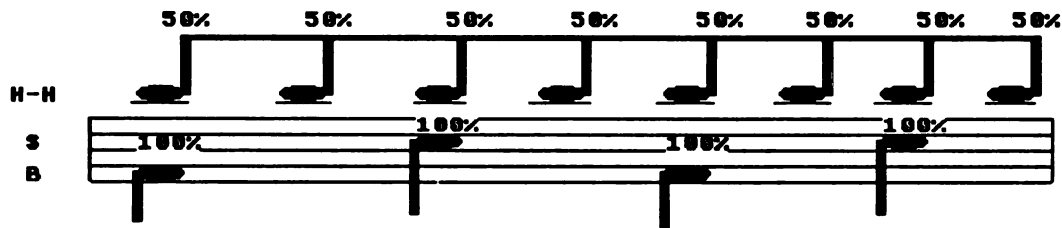


Figure 4. Sample Pattern With 1:1 High-Hat Duration Ratio

By experimentation it was found that skewing the 1:1 ratio by around 13%, making it approximately 3:2, produced a shuffle rhythm, such as is found on Jimmy Reed's, "Bright Lights, Big City" and many other blues and jazz pieces. In this case, the first eighth note occupies 63% of the beat and the second occupies 37% (see Figure 5).

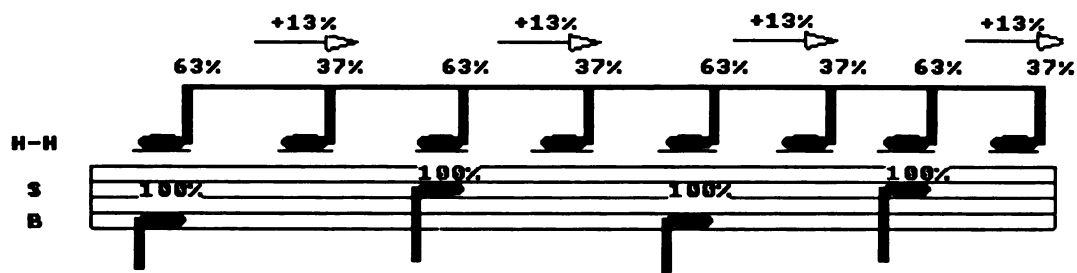


Figure 5. Sample Pattern With 13 Percent Skew in High-Hat Durations

An interesting problem arises when trying to notate this shuffle rhythm. Indicating an exact 3:2 relationship is quite cumbersome. It can be done by dividing the beat into five equal parts or quintuplets. The first three quintuplets and the last two can then be tied together to arrive at the proper duration and proportion (see Figure 6).

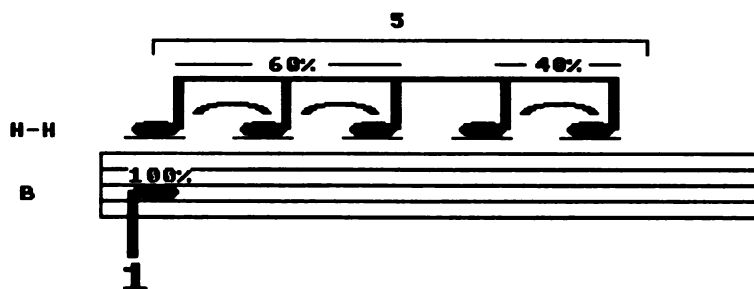


Figure 6. Using Tied Quintuplets to Represent a 3:2 Relationship

An alternative notation is the dotted eighth/sixteenth note combination, or, a 3:1 ratio. This is the clearest way to *imply* the shuffle rhythm, but if it is performed too literally, as a MIDI sequencer would, it sounds too choppy for an authentic shuffle (see Figure 7).

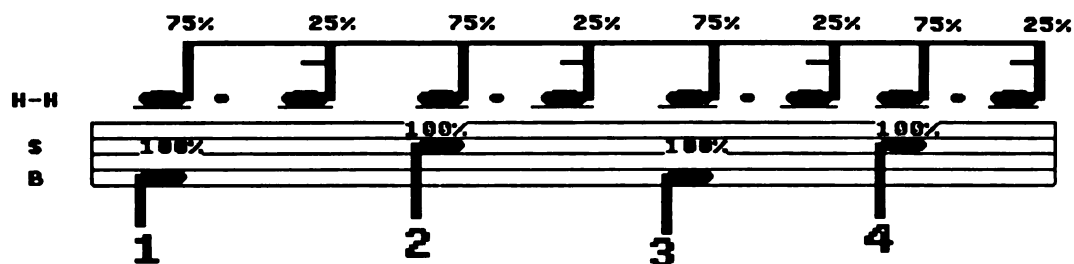


Figure 7. Approximating a 3:2 Relationship With Dotted Eighth-Sixteenth Combinations

Getting a computer to produce the shuffle rhythm requires some individual note editing. At 240 PPQ, the first eighth note is edited to a duration of 150 and the second becomes 90. Changing only the high-hat note ratios is enough to create the shuffle rhythm, the bass and snare drum notes may remain unchanged.

The 13% difference that accounts for the shuffle rhythm would fall under the heading of systematic variance, where it is notated as series of dotted eighth-sixteenth combinations. The other style of notation, using tied quintuplets, is cumbersome and requires more mental effort from the drummer translating it into sound. And even then, the 13% shift is not exactly a 3:2 ratio.

The shuffle rhythm does not fit neatly into the traditional system of music notation. Hence, it is heavily dependent on extra-notational time valuation.

It is possible for the music synthesist to build up knowledge of the manner of systematic variance typically applied to various common structures and to use it as a starting point when trying to make synthesized music sound more natural. This knowledge of systematic variance is an area of fruitful study for music synthesists, because it can enable them to make their creations more realistic, while enjoying the freedom and economy offered by the technology.

Trying to establish a body of knowledge of typical systematic variance applied to common musical structures for application to synthesized music requires a different method than that used by musicologists in "pure research." It must be directed toward the goal of *applicable knowledge*. Thus, the equipment chosen must be typical of a large number of users. And the pieces of music used for measurement must be representative of a particular style, preferably one which has wide popularity. This gives any reliable results a wider currency.

In the present investigation, two alternate methods of registration have been used. The first method uses audio tape and grease pencil marks to measure note duration; the second method uses a MIDI sequencer to measure note duration and intensity (velocity). The musical genre explored is Rock and Roll and the instrument is the drums.

TAPE AND PENCIL REGISTRATION OF MUSICAL RHYTHM

Method

This method used grease pencil marks on the backing of audio tape to indicate the start times of notes in performances recorded on the tape. It proved a simple, fairly accurate way to look for systematic variance in rock drum rhythms. Several musical selections were recorded on a half-track stereo reel-to-reel tape deck at a speed of 15 inches-per-second. Then, the tape was played back manually, rocking the reels back and forth slowly while in contact with the playback head to determine the start points of the bass drum and snare drum hits. Bass drum hits were marked with a blue pencil, snare drum hits with white.

A transcription into musical notation for the piece was taken down on paper. Next, the distances between the bass and snare drum hits were measured to a thirty-second of an inch and written on the paper above the notes. Then, the distances and the total length of the piece were calculated in terms of milliseconds. This provided the individual note lengths in milliseconds and the tempo of the selection. Finally, the millisecond values were transformed into PPQ values, based on the 240 PPQ resolution of the sequencer used in the investigation.

The accuracy of this method was estimated as follows. At 15 ips, one inch represents $1/15$ of a second, or .067 seconds. Dividing one inch into 32 parts yields the value in seconds of $1/32$ of an inch: $.067/32$ equals .002. Thus, one thirty-second of an inch represents 2 msec, at 15 ips. Given that the grease pencil mark itself can be as much as one sixteenth of an inch wide, and that determining the exact beginning of a note attack sometimes involves a little guess-work, it is safe to say this method is at least accurate within 4 to 10 msec.

The musical examples taken for this investigation included other instruments playing at the same time as the drums. Among them were, "I'm Ready," from Humble Pie's, "Rockin' The Fillmore" album; Robin Trower's, "Fool and Me," from his, "Bridge of Sighs" album; and, "Good Lovin' Gone Bad," by Bad Company, off "Straight Shooter." An exception was "I'm Ready," from which a solo drum introduction was taken. Short sections of the other pieces were taken that represented the basic beat of the tune.

To verify the accuracy of the registrations the note values in PPQ were entered into the sequencer program and played back. In a few cases, slight editing of start times was necessary to make the piece sound like the recording.

Once this was done, the sequence would have a feel similar to that of the original.

Results

Though this method is very tedious and gives sometimes spurious results, it provided some interesting insights. When examining the registrations, many deviations from rational-mechanical note duration were found. Some of these differences were small, but many were quite large.

The deviations from the rational-mechanical norm were analyzed on two levels. First, as deviations from "normal" note duration producing changes in tempo between measurement nodes. Deviations on this level are changes in *micro tempo*. The measurement nodes are points in the registration where it is convenient to measure the duration of a note or series of notes. For instance, in a situation where there is a quarter note and two eighth notes in series, the two eighth note durations would be added together. There would then be two nodes: one consisting of the quarter note and another made up of the two eighth notes. This procedure makes calculating tempo changes easier, since tempo is defined as the number of quarter notes per minute. This grouping of note durations into nodes is also necessary due to the difficulty of determining the durations of rests in the registration. Since the drum sounds are mixed with other instrument sounds, it is usually impossible to determine the precise point where the sound of a given drum tone ceases and where the rest begins. Thus, a rest succeeding a note must be added to that note's duration for the purpose of calculating deviation from the rational-mechanical norm for the duration of the note/rest structure.

An example will best illustrate the calculation of micro tempo. In the registration of "Good Lovin' Gone Bad," the first quarter note has a duration of 450 msec. The total duration of the sample is 10.9 sec. The sample contains six 4/4 measures, or, a total of 24 quarter notes. Dividing 10900 by 24 gives the duration in milliseconds of one rational-mechanical beat for the sample. The duration of one beat equals the duration of a quarter note, by definition. In this case the duration is 454 msec. By multiplying or dividing 454, the rational-mechanical durations of other note lengths can be found. For instance, $2 * 454 = 909$ = the duration of a half note, etc. The rational-mechanical tempo must also be calculated. This is done by dividing 60 by .454. This yields, 131, the number of 454 millisecond beats (quarter notes) that fit into a second. Once the rational-mechanical tempo and note durations have been established, the deviations can be calculated. For this first quarter note of the sample, the deviation is $450/454$, or, .9911894.

One could say it has about ninety-nine percent of the rational-mechanical duration. As for its effect on micro tempo, it speeds it up by around one percent. To reflect this fact the rational-mechanical tempo of 131 is multiplied by $1/.9911894$, yielding 132. Thus, micro tempo across this quarter note is 132.

A computer program was written in Advanced Basic to carry out these calculations based on the actual note durations from the registration. The ASCII output of this program, including the deviation per node value in terms of beats per minute, was then transferred to Harvard Graphics (TM) so that the micro tempo changes could be graphed.

The output of Harvard Graphics (TM) is shown below. A direct comparison of the three registrations is not possible, since the number of quarter notes contained in the respective nodes is not always equal. Even so, more general comparisons can be made of the samples, such as the magnitude and direction of the variance at the beginning, middle and end of each, the overall magnitude of the variance, etc.

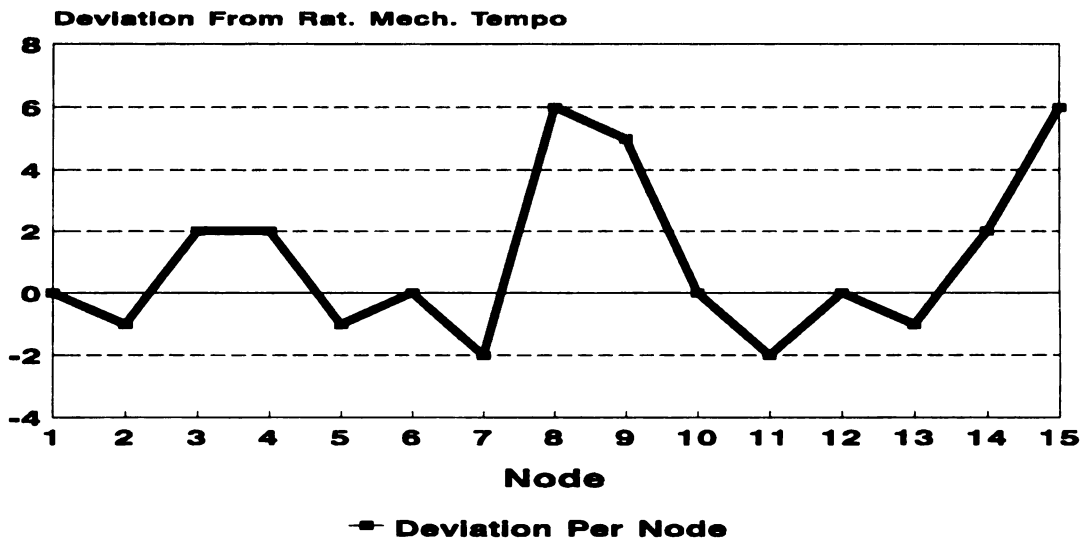


Figure 8. Micro Tempo Map for "I'm Ready"

The segment from "I'm Ready" begins on the zero line, that is, at the rational-mechanical tempo. It has a large micro tempo increase in the middle (to +6 beats per minute) that is compensated quickly, and another increase to +6 bpm at the end. There are four points in the segment that lie on the zero line.

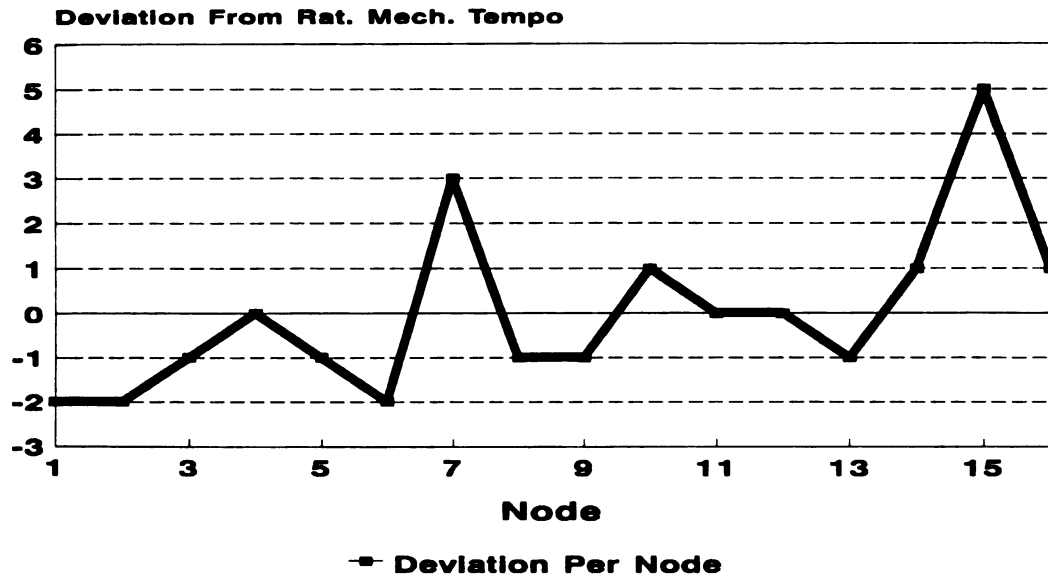


Figure 9. Micro Tempo Map for "Fool and Me"

The segment from "Fool and Me" has a slow start. In other words, the micro tempo begins below the zero line and remains there for a few consecutive nodes. The fastest points occur in the middle (+3 bpm) and at the end (+5 bpm), both of these points are followed by immediate compensation. There are three points in this segment that lie on the zero line.

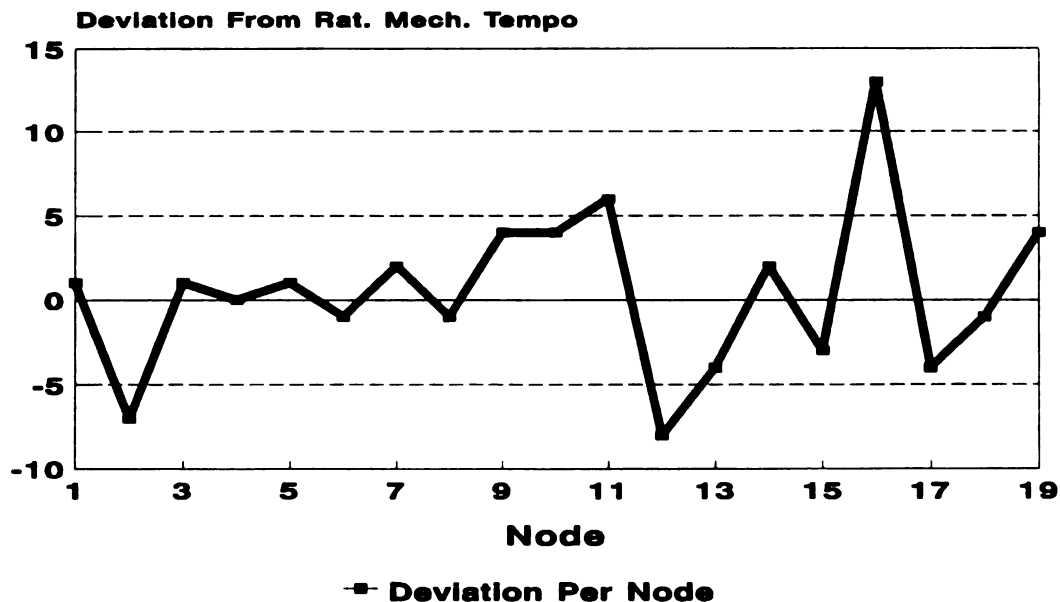


Figure 10. Micro Tempo Map for "Good Lovin' Gone Bad"

The segment from "Good Lovin' Gone Bad" starts slightly above the rational-mechanical tempo, but has a large slow down (to -7 bpm) at the second node. This is followed by a stretch where the micro tempo hovers around the zero line. About three quarters through the segment, there is a speed up (to +6 bpm) followed by a sharp slow down (to -8 bpm). Near the end is the fastest point (+13 bpm), followed by immediate compensation. Then there is a rise to a micro tempo 4 bpm faster than the rational-mechanical norm as the segment ends. There is only one point in this segment lying on the zero line.

On another level, the deviations can be viewed as a long term trend, in other words, as constituting a gradual speeding up or slowing down of the tempo from the beginning to the end of the sample. That there are any such long term tempo changes in the sample segments is not evident looking at them in terms of micro tempo. The wide deviations from rational-mechanical tempo appear, more or less, to be compensated in every case. However, a linear regression line derived from the actual note durations, if it proves reliable, is a more accurate way to assess if there are any long term (macro) tempo changes happening in the segments.

In order to generate the x and y coordinates necessary to calculate the regression line for a given registration, something called *progress scores* were generated. The progress scores are derived from performing cumulative addition on the actual and rational-mechanical duration values. This yields a set of x and y coordinates that produces a line showing the deviation from rational mechanical tempo as the segment progresses.

For instance, in the above example, the first note of "Good Lovin' Gone Bad" has a duration of 450. It's rational-mechanical norm is 454. The coordinates of the first point, then, are (454, 450). The duration of the next note in the registration is 481. It's rational-mechanical counterpart has a duration of 454, also. The coordinates of the next point would be $454 + 454 = 908$, for the rational-mechanical norm on the x axis, and $454 + 481 = 931$, for the y component, making the ordered pair, (908, 931). The process continues on until the last coordinates are generated, which are in fact the actual and rational-mechanical total durations for the registration.

As such, these coordinates are simply another way of indicating micro tempo: they correspond to the microtempic behavior described above for this particular segment. However, when these points are smoothed into a straight line by the regression formula, the slope and intercept of the regression line tell of the long term tempo changes within

the registration; they can be compared to the rational-mechanical slope and intercept of 1 and 0.

The formula for a line, $y = a + bx$, states that a given x value (in this case the rational-mechanical norm) is multiplied by a constant, b (the slope), and added to a to obtain the corresponding y value. In the context of the progress scores and the line derived from them, a slope of less than 1 means the value of bx will be smaller than x . A smaller bx represents a shorter duration than the rational-mechanical norm, and, hence, a faster tempo. Of course, the value of bx is modified by a (the intercept), and a bx smaller than x (faster relative tempo) can be negated or over-ridden by a large a . However, as the segment progresses the x values rise up into the thousands of milliseconds. At this point an intercept of, say, 20, has much less effect on the obtained y value than a slope of, say, .98.

Consider the following example from "Good Lovin' Gone Bad," using the above slope and intercept of .98 and 20. The rational-mechanical value of the first note is 454. $454 * .98 = 445$; $445 + 20 = 465$. The difference between 465 and 454 is 11. The value of 465 indicates that tempo is slower than normal across this first note, according to the regression line's equation. In this case, the less than 1 slope is over-ridden by the intercept.

Now, a point (node) toward the end of the registration is chosen. At node 15, the rational-mechanical value is 8630. $8630 * .98 = 8457$; $8457 + 20 = 8477$. The difference between 8630 and 8457 is 153. The value generated by the regression equation is less than the rational-mechanical norm. Therefore, tempo is faster than normal at this node.

From the preceding it is clear that the slope and intercept together describe the deviations from normal tempo in the macro sense. They indicate how fast or slow the selection begins relative to the rational-mechanical tempo, and what gradual tempo increases or decreases occur in the long run.

Correlating the rational-mechanical note durations with the actual durations, the strength of the relationship between the actual performance and its abstract form, the rational-mechanical norm, can be assessed. The performance must be a variation of its abstract form, because the rational-mechanical values are computed post hoc, from a transcription of the performance. Therefore, the correlation between the actual and rational-mechanical note durations is expected to be high. During this part of the assessment process, the progress scores were not used, because their artificial linearity (created by the cumulative addition) would overstate the strength of the

relationship between the performance and its rational-mechanical counterpart.

The regression line slopes and intercepts for the three registrations are given below, along with the correlation coefficients. These were computed using using SPSS/PC+ (TM).

Table 1. Tape and Pencil: Correlations, Actual and Rational-Mechanical Durations.

<u>Registration</u>	<u>Act./Rat.</u>
Ready	.9999
Good	.9967
Fool	.9993

Table 2. Tape and Pencil: Slopes and Intercepts.

<u>Registration</u>	<u>Slope</u>	<u>Intercept</u>
Ready	.97853	4.95
Good	.99907	15.00
Fool	.99998	21.35

The high correlations between the actual and rational-mechanical durations suggest the respective linear models will be accurate in predicting the actual progress scores.

The accuracy of a regression equation can be checked by examining its residuals. To obtain the residuals, predicted values are generated from the rational-mechanical progress scores and the regression equation. Subtracting the predicted from the actual progress scores produces the residuals, which are simply the difference between the two at each node. A residual of 0 means that the equation predicted perfectly the actual progress score at a particular node; consistently high residuals mean that the equation is a poor description of macro tempo changes in the registration. High residuals may also mean that the linear model is unfit for describing macro tempo in that case.

Below is a residual map for "Good Lovin' Gone Bad". The line marked "0" represents the regression line described by the equation. The crooked line snaking around it is the

residual map. Were the equation perfect in predicting actual progress scores, the residual map would be equal with the zero line. In this case, there are mixed results. The equation is as close as 2 msec in its prediction, and as far off as 33 msec.

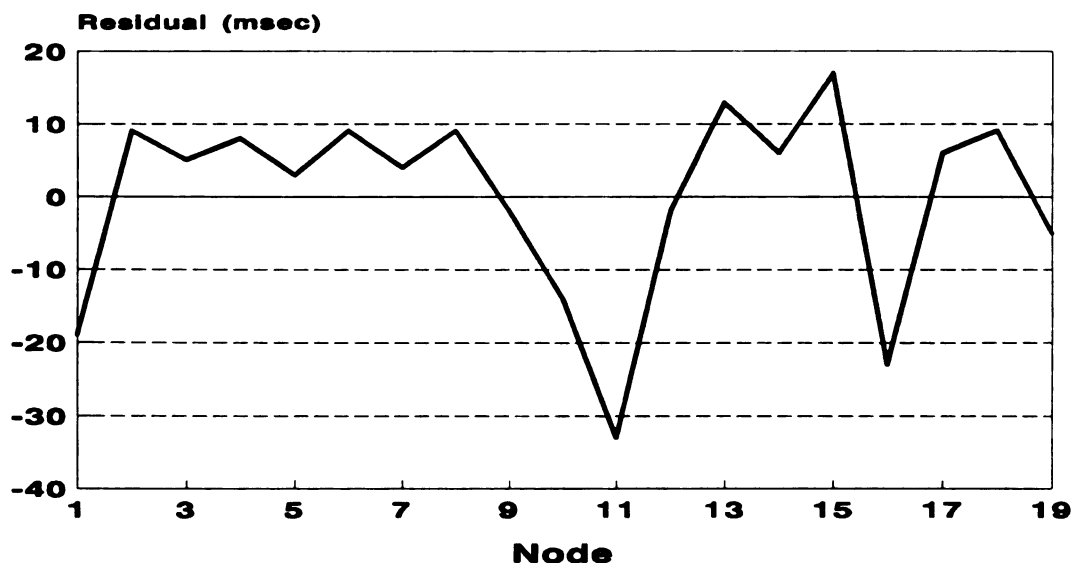


Figure 11. Residual Map for "Good Lovin' Gone Bad"

Though the residual map appears to fluctuate wildly at points, the magnitude of the worst predictions must be kept in perspective. The largest residual equals 33 msec. The closest note value to 33 msec at this tempo is a sixty-fourth note, at 28 msec. This means that, worst case, the equation is off by roughly the duration of a sixty-fourth note, a remarkable degree of accuracy.

It is acceptable to compare the three registrations in terms of their regression lines. "I'm Ready" has the flattest slope, and the smallest intercept value. It is clear "I'm Ready" will have the greatest tempo increase in the long run because of its slope. "Good Lovin' Gone Bad" has a positive intercept and the second flattest slope of the three. This means it has somewhat of a slow start, but speeds up over its course. "Fool and Me" has the largest intercept value and the steepest slope of the three. This means it starts up the slowest and speeds up the least.

At this point, the progress scores are helpful in interpreting the macro tempo changes the regression lines are describing. Table 3 shows the actual and rational-mechanical total durations for the three registrations.

Table 3. Actual and Rational-Mechanical Total Durations for the Three Tape and Pencil Registrations.*

<u>Registr.</u>	<u>Act. Total</u>	<u>Rat. Mech. Total</u>	<u>Difference</u>
Ready	11,846	12,116	-270
Good	10,900	10,900	0
Fool	7940	7936	+4

* Minus sign means the actual duration was short of the rational-mechanical norm; plus sign means actual duration is in excess of the rational-mechanical norm.

Here it is clear that "I'm Ready" has a discernible tempo increase: its actual duration is 270 msec shorter than the rational-mechanical norm. "Good Lovin' Gone Bad" has an actual duration equal to the norm, meaning that its slow start is perfectly made up by its tempo increase. For "Fool and Me," the slow start is not completely made up in the tempo increase: its actual duration is a bit longer than the rational-mechanical norm.

While the difference for "I'm Ready" appears substantial, it is important to keep it in perspective. The duration of an eighth note at the rational-mechanical tempo for "I'm Ready" (which is 132) is 227 msec. The 270 msec difference, then, makes the actual duration short by approximately an eighth note. In other words, the loss in duration amounts to about an eighth of a measure by the end of the segment. The difference for "Fool and Me" is, of course, inconsequential, since it falls within the estimated margin of error for the Tape and Pencil method, which is 4 to 10 msec.

Discussion

The short length of the sections--from about 8 to 12 seconds--limits somewhat the amount of knowledge to be gleaned from these analyses. It is more desirable to have the whole song registered for analysis, in order to get a clear idea of what a drummer is doing in a song, and where relative to its global structure.

Still, it is useful at this stage to have registered some rock drum performances and worked out a meaningful analysis scheme. The tempo increase found in each of the performances, whether compensated or not by a slow start is curious because only "I'm Ready" was taken from the beginning of the song. One can understand a slow start at the very beginning of a song as a way to "work up to" the basic tempo of the piece. However, the "Good Lovin'" and "Fool and Me" samples were each taken at the start of a

verse near the beginning of the piece. Moreover, if the tempo increase implied by the slope were continued indefinitely, it must certainly become noticeable at some point. It seems likely that if these gradual tempo increases are a common characteristic of rock drumming, there is some point where the increase reverts to a slower tempo, either gradually or instantaneously. This may happen at phrase points (the endings of verses and choruses), where roundings or fills are often inserted in place of the repetitive beat.

The Tape and Pencil results also suggest why machine generated drum patterns are so readily discernible to listeners. A drum box, even if it can introduce systematic variance to affect the micro tempo, will not routinely add a macro tempo change on top of it. Even if it did, there would still be the question of where to revert to a slower tempo, lest the pattern's tempo increase indefinitely.

Finally, these short registrations are valuable because they inform on how to construct the studio experiment, where a human drummer will be asked to play certain patterns into a PC-based MIDI sequencer. Hypotheses may be generated at this point which the studio experiment can be designed to test.

The tape and pencil method, *per se*, seems most useful where the musician wants to study the rhythmic character of a specific piece of music or drummer. However, where the drum sounds are mixed with other instrument sounds, the accuracy of the registration declines. This is because the bass drum and cymbal attacks are many times masked by other instruments. Nevertheless, the snare drum hits are usually very distinct. These snare drum hits alone can provide useful data on micro tempo and macro tempo. Also, bass drum hits tend to be more pronounced at points of accent; therefore, the method may be useful for studying how different drummers accent various common rock drum structures (i.e., beginnings of measures in similar patterns).

Furthermore, a multitrack recording of a given piece of music would allow the analyst to customize the mix of the piece in functional ways. If the start times of the cymbal hits are needed, a special mix can be made that emphasizes the cymbals at the expense of the other instruments, even leaving some of the others out altogether. Also, the special mix could be dubbed at 30 ips, spreading notes farther out on the tape physically and making the registration more accurate.

This option is open to a researcher with access to a band and a multitrack recording facility, or to a musician who possesses or has access to some multitrack recordings.

Unfortunately, it is probably impossible to obtain multitrack copies of commercial music releases.

While it allows the musician to make use of a convenience sample (i.e., his own record collection or multitrack recordings) in studying systematic variance, the tape and pencil method has some critical limitations, such as, the difficulty of finding all the notes attacks in the mix, and the fact that note intensity cannot be registered.

Its main drawback, however, is the sheer tedium of manually playing back the tape, marking the attacks, and measuring the distance between them. This process is so time consuming and cumbersome as to place serious limits on the amount of data that can be collected with this method. The next method to be described overcomes this by using the power of a personal computer to automate parts of the data collection process, significantly reducing the time interval between data collection and statistical analysis.

MIDI SEQUENCER REGISTRATION OF MUSICAL RHYTHM

This method uses a human drummer performing rhythms into a MIDI sequencer. To begin, some rock and roll drum rhythms were selected from several "rhythm boxes." The boxes used were a Yamaha *DD5 Digital Drums* unit, a Yamaha *PSR 36* electronic keyboard and the Roland R-8, *Human Rhythm Composer*. The rhythms are designated on the boxes with names such as, "Rock and Roll," "Slow Rock," "Hard Rock," "Rock 1," etc. These rhythms were transcribed and transferred to the MIDI sequencer. There were, then, several computer files containing the various rock drum rhythms. An analog tape of the sequences would be made for a drummer to study. The drummer would be instructed to learn to play them exactly or as closely as naturally possible, but in his own style. He would be given about a week to practice the material.

As it turned out, two sessions were required in order to obtain a satisfactory amount of data. The second session was necessary because of certain technical complications in the first session. Each session employed a different drummer and an somewhat different set-up. Each will be described separately below.

For the first session, the set-up was as follows. The drummer brought his acoustic drums to the twenty-four track recording facility at Michigan State University. A small set was used, consisting of bass drum, snare drum and high-hat. The drums and cymbals were miked and each microphone signal was fed to a Keypex expander. The Keypex's were used to gate the microphone signals from the individual drums, to combat the interference resulting from how close the drums and microphones were placed to each other. False triggering would result if, for instance, the snare drum's sound was picked up by the high-hat microphone. By careful setting of the gate threshold for the high-hat microphone, it was thought possible to reject the snare drum sound while picking up the high-hat sound reliably.

From the expander the microphone signals were routed to a device called a *MIDI KITTY*. The *MIDI KITTY* converted the processed microphone signals into MIDI note data. The MIDI OUT jack of the *MIDI KITTY* was connected to the MIDI IN of the sequencer and the MIDI OUT of the computer sequencer was connected to a Proteus/1 XR synthesizer (by E-Mu Systems, Inc.) containing the sampled drum sounds. The computer running the sequencer was a Zenith Z-150, IBM XT compatible.

Out of the synthesizer, the audio signals for the synthesized bass drum, snare drum and high-hat sounds were routed to an Otari twenty-four track tape deck. Coming from the tape deck and into the computer's MIDI interface card (a

Music Quest MXQ16S) was a SMPTE time code signal which had been striped to track 24 of the tape.

There were two reasons why this set-up was chosen. First, because a set of MIDI drum pad controllers was not available at that time. Second, because it was thought that allowing the drummer to play his own drum set would produce more representative results.

As a part of setting up, it was necessary to do some checks of the equipment performance. Specifically, it was important to estimate the amount of MIDI and processing delay. To do this, the acoustic signal from the bass drum microphone was routed to the left channel of a half-track stereo tape deck, running at 30 ips. The output from the tone generator, that is, the synthesized drum tone, was routed to the right channel of the tape deck. A recording was then made of the drummer hitting the drum. A registration of the recording using the tape and pencil method yielded an estimate of the system delay from hit to sound of about 7 msec. At 30 ips, one thirty-second of an inch equals one msec.

Once the set-up was ready, the Drummer A was called in to do some playing. He was instructed to make five passes at the rhythms he had been given, playing each for about sixteen measures, with a drum fill inserted in the eighth measure. He was allowed to hear the first few takes back to check if they were acceptable. Finding they were, he performed the rest without listening to them back, occasionally requesting to do a take over.¹¹

Also, one version of each beat was recorded to the multi-track tape. The recordings consisted of six tracks. Three tracks were the acoustic sounds of the three percussion instruments (bass drum, snare drum and high-hat) and three were used for the synthesized percussion sounds.

A one measure count was given to the drummer from the sequencer. This count came down the MIDI cable and was routed to the drummer's headphones. The drummer was instructed to give a one measure count of his own (on the bass drum) as an extension of the sequencer's, prior to playing the prescribed pattern. From this it could be determined how well the drummer perceived the tempo from the sequencer's count.

11. This procedure is similar to one used by Gabrielsson. It is supposed to ensure the performances are musically acceptable, and, thus, representative. Ideally, the drummer should have listened to and judged every take, but it seems he could tell without listening back if the performance contained any mistakes, so he was allowed to continue in that way.

A variation on this scheme was tried. The drummer played one of the selected beats, only this time with a guitarist accompanying him. The guitar performance was recorded to the tape deck, along with the various percussion sounds as described above. Also, the MIDI data was recorded by the sequencer.

The set-up for the second session was designed to address certain technical difficulties encountered in the first. While the MIDI KITTY is a versatile means of converting acoustic drum sounds to MIDI note information, it is apparently not designed to use a microphone signal as an input. The manual makes no mention of using microphones as input devices, rather, it assumes the use of drum triggers. A drum trigger is mechanically coupled to a drum or cymbal with an adhesive. It transforms mechanical vibration into an electrical signal for processing by the MIDI KITTY. Since it does not operate on the principle of sound pressure, interference between the various parts of the drum kit is reduced, because the triggers are easier to mechanically isolate than the microphones are to sonically isolate. Also, the sustain portion of a trigger's attack envelope is shorter than a microphone's. Thus, though the MIDI KITTY is equipped with sophisticated facilities for eliminating cross-triggering problems (where a hit on one drum triggers another's sound in addition to its own), it seems that it will work best with drum triggers, and less satisfactorily with microphones, due to the different natures of these two types of transducers.

The problem of cross-triggering was handled with both the Keypex's and the MIDI KITTY's controls. The greatest difficulty was between the high-hat and the snare drum. The snare drum is a louder instrument than the closed high-hat. The high-hat was placed in the usual position, to the left and within ten inches or so of the snare drum. In order to isolate the high-hat mike from the snare drum hits, it was necessary to raise the gate threshold for the high-hat mike so that the high-hat hit just opened the gate and triggered the high-hat sound through the MIDI KITTY.

Doing this, however, created a new problem. Using the gates in this way has the effect of reducing the drummer's available dynamic range. If he plays soft enough, the drum's gate will not open and the sound will not trigger the MIDI KITTY and will not be registered by the sequencer. In extreme situations, such as with the snare and high-hat, the dynamic range can be so crushed as to leave the drummer very little flexibility in performance dynamics, which are important in accenting. In the first session, cross-triggering and the dynamic range problem made it necessary to adjust the Keypexes and the MIDI KITTY constantly. This was a time consuming task that wearied the drummer and

reduced the amount of data collected in the session. Thus, the second session.

Shortly after the first session, it was learned that another drummer was available who had a set of electronic drum pads. Another date was scheduled with a slightly different set-up.

The Drummer B brought his set of drum pads and his own triggering device, a unit by Simmons, called a TMI Trigger/MIDI Interface. The MIDI KITTY was on hand if the TMI unit proved to have excessive MIDI delay. This arrangement was much more convenient because the drummer knew his equipment well and configured it quickly and just as desired. Six pads were used to trigger the sounds of the high-hat, snare drum, bass drum, tom tom, ride cymbal and crash cymbal. They were situated in the general positions that the real percussion instruments would be in a conventional drum kit. The MIDI OUT of the TMI unit was fed to the sequencer and the sequencer output to the Proteus/1. A single output of the synthesizer was used this time to transmit all the drum sounds to a single track on the tape deck.

Of the six drum beats selected prior to the first session, the first drummer performed three. The goal for the second session was to get data on the other three drum patterns, and to get additional data on guitar accompanied drum performances.

As in the first session, the drummer was given a four beat count through his headphones and instructed to continue that count for four beats on his bass drum as an intro to his performance of a given beat. He listened back to the first few performances, after which he chose to continue on without listening to each performance. The Drummer B was required to do only two repetitions of the patterns, because five seemed to fatigue the Drummer A. Drummer B was paid for his work.

The MIDI delay was measured by placing a microphone in front of the snare drum pad and having the drummer hit the pad with a stick while the microphone and the synthesized signal were recorded on the two track machine at 30 ips. The delay was measured using the tape and pencil method and found to be about 11 msec. This was longer than for the MIDI KITTY. The drummer seemed to make up for this delay by playing slightly ahead of the beat. Since the beats, as heard from the synthesized drum tones, sounded rhythmically acceptable, it was decided to stay with the TMI unit, though it was slower, because it was working well otherwise and the drummer was familiar with its performance.

Pretest of MIDI Sequencer Registration Method

Before going into the studio, it was thought best to pretest registration method two. The purpose was to refine the methods of analysis established in the Tape and Pencil Method, make judgments on the number of repetitions required for each pattern, and look for any additional data on rhythm performance with which to refine the studio experiment.

The Yamaha DD5 has four velocity sensitive drum pads on it and a MIDI OUT jack. While the cramped orientation of the pads is not ideal for natural playing, yet the DD5 sufficed as a drum controller in the pretest. Its MIDI OUT jack was connected to the MIDI IN of the PC sequencer so that the MIDI controller data could be recorded. The sequencer was set to send a four beat count via MIDI to the drummer's headphones after record was activated on the sequencer. The count-down procedure prior to beginning the pattern was followed as outlined above. The tempo was set at 93; the author acted as drummer for the pretest.

Ten versions of a seven measure rock drum pattern were performed into the sequencer, five on one day, five on the next. The ten versions were recorded as individual tracks in a single sequence. Each track could be listened to independent of the others. Once the ten versions were judged "acceptable," that is, free of mistakes and "average sounding," the analysis began.

The pattern (Figure 12) contained five notes and was one measure long. It was thus repeated seven times in each version.



Figure 12. Pretest Drum Pattern Notation.

The first step was to take down the duration of each note in the ten versions. This was done by measuring the distance in PPQ between the start times of consecutive notes. Once this was finished, the duration data was formatted into a data file for use with the SPSS/PC+ statistical package.

The first step in the analysis was to run some descriptive statistics on the whole sample (ten versions). For each of the seven measures, SPSS calculated a grand mean, median and mode for each of the five notes in the pattern. The next step was to perform t-tests on the note means to see if there were any intra-measure significant differences in the durations of consecutive notes with equal notated duration values.

A strong significant difference was found for the durations of the three eighth notes at the end of the pattern. Table 4 shows the note durations and the observed significance levels.

Table 4. Pretest: T-Tests of the Three Eighth Notes.*

<u>Meas.</u>	<u>Note 3 Dur.</u>	<u>Note 4 Dur.</u>	<u>3-4 Sig.</u>	<u>Note 5 Dur.</u>	<u>4-5 Sig.</u>
1	123	111	.000	123	.009
2	126	114	.000	117	.013
3	120	116	.000	114	.044
4	122	109	.002	115	.227+
5	120	108	.000	113	.034
6	120	112	.000	115	.027
7	116	109	.000	116	.017

*The rational-mechanical duration of the eighth-note is 120. "+" indicates the difference did not reach at least the .05 significance level. Note durations given are the averages across the 10 versions.

In every case except one (.002 in measure five) the difference between the means of the third and fourth notes was significant at the .000 level. Likewise, the relationship between the means of the fourth and fifth notes was significant to at least the .05 level in all but one case (.227 in measure five). This indicates there is systematic variance being applied to the last three notes of the pattern.

As another test, the deviation from the rational-mechanical norm for every note in each of the ten versions was calculated. A factor analysis was then performed on these ten sets of deviations. Three factors were found to account for 68.6% of the variance. All but two of the ten versions belonged clearly to only one factor. The factors represent different characteristic ways of performing the pattern. Table 5 shows the factor loadings for the ten versions.

Table 5. Rotated Factor Loadings for the 10 Versions.

<u>Version</u>	<u>Factor 1</u>	<u>Factor 2</u>	<u>Factor 3</u>
1	.01981	.18430	.84418*
2	.68387*	.09946	.37659
3	.27976	.82069*	.27379
4	.00349	.88710*	.20956
5	.65766*	.04710	.27235
6	.61136*	.46683	.39965
7	.76578*	.20121	-.10888
8	.23117	.14158	.69963*
9	.52633	.61534*	-.06594
10	.75046*	.37656	.10144

*A loading of greater than .60, indicated by an asterisk following the number, indicates the version conforms mainly to the respective factor.

Using a value of .65 as the cut-off point, eight of the versions can be assigned exclusively to one factor. The objectionable versions are number nine, with a .526 loading for factor 1 and .615 for factor 2, and number six, with loadings of .611, .466 and .399 for factors one through three, respectively. Versions nine and six are perhaps hybrids, where a part of each conforms to a different factor.¹²

Once the versions are assigned to their respective factors, the note durations of the versions belonging to a particular factor can be averaged to obtain the "characteristic version" represented by that factor. Descriptive statistics, such as the mean and deviation from rational-mechanical norm, can be used to explore differences between the different characteristic versions. However, inferential statistics, such as the t-test, are not very useful in this case, since the low sample size (n=4 in the case of factor 1, and n=2 for factors 2 and 3) make even large differences between note duration means appear insignificant. Nevertheless, something can be said about the similarities and differences between the three characteristic versions (CV's).

On the intra-measure level, in one-hundred percent of the cases, the long-short relationship between notes three and four is maintained. The short-long relationship between notes four and five is maintained eighty-six percent of the time in CV1, forty-three percent of the time in CV2, and seventy-one percent of the time in CV3.

12. A .64 cut-off is used by Gabrielsson (1980) to interpret factor analyses of the rhythms in one of his experiments, elsewhere, he uses .70. Using .64, eight of the ten versions can be assigned to a factor; using .70, five of the ten can be assigned.

Another means of comparing the three characteristic versions is in terms of their *micro tempo maps*, as done in the Tape and Pencil registrations. Tempo changes were computed at three nodes in each measure: at the first quarter note, across the dotted quarter-eighth combination, and across the last two eighth notes.

The method of calculating the micro tempo deviations from rational-mechanical tempo for the sequencer registrations was very similar to that used for the Tape and Pencil registrations. The main difference is that the sequencer expresses note durations in terms of PPQ, instead of milliseconds. The computations were performed by the same BASICA program used above, utilizing an option that considers the 240 PPQ beat resolution of the sequencer used in the investigation.

As above, the deviations were input to Harvard Graphics for a visual representation. The results are given on the following page. They demonstrate the differences between the three CV's. It is interesting the three have very similar ranges of deviation. The range for CV's 2 and 3 is -2 to +4 bpm; the range for CV1 is -3 to +4 bpm.

The final comparison of the three characteristic versions is in terms of macro tempo, or, their respective regression equations. The results are given below, in tables 6 - 8.

Table 6. Pretest: Correlations, Actual and Rational-Mechanical Durations.*

<u>Registr.</u>	<u>Act./Rat.</u>	<u>Sig.</u>
CV1	.9990	.000
CV2	.9991	.000
CV3	.9989	.000

* "Act./Rat. is the correlation between the actual and rational-mechanical note durations.

Table 7. Pretest: Slopes and Intercepts.

<u>Registr.</u>	<u>Slope</u>	<u>Intercept</u>
CV1	.98711	26.42
CV2	.99400	21.55
CV3	.99666	11.99

Table 8. Pretest: Actual and Rational-Mechanical Total Durations.*

<u>Registr.</u>	<u>Act. Total</u>	<u>Rat. Mech. Total</u>
CV1	6640	6720
CV2	6688	6720
CV3	6711	6720

* Values are in PPQ.

By examining the micro tempo maps and the regression equations, it becomes clear that the three CV's differ significantly. CV1, the dominant performance style, or, factor, has the slowest start and the fastest increase. CV2 has the next slowest start and the next fastest increase, and CV3 has the least slowest start and the least increase in tempo over its duration. As for the overall increase from rational-mechanical tempo found in each of the three

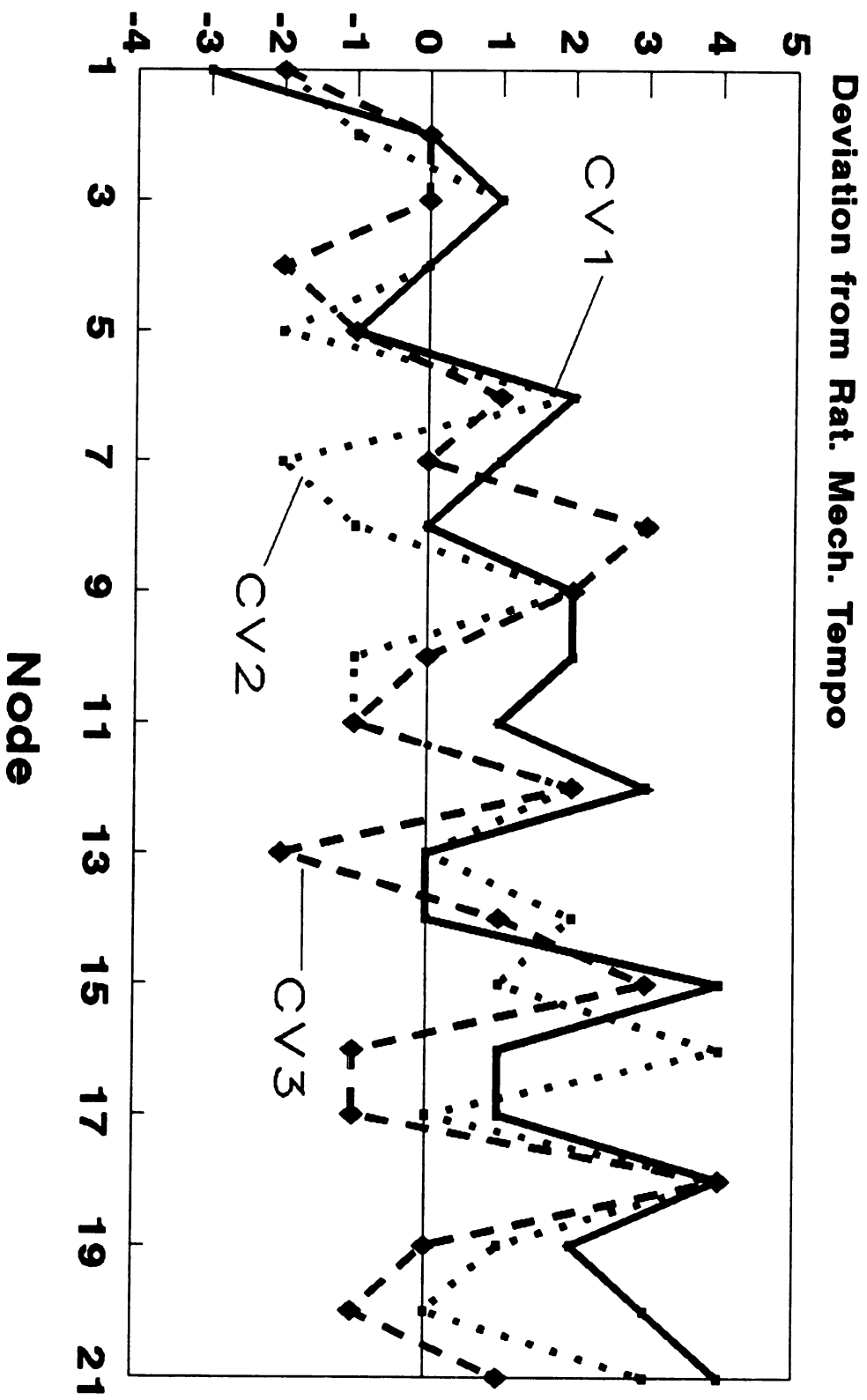


Figure 13. Microtempo Maps for Pretest

CV's, CV1 and CV2 both have actual durations quite short of the rational-mechanical norm. CV3 has a total duration only 9 PPQ short of the rational-mechanical norm. For the sake of comparison with the Tape and Pencil registrations, it is useful to convert the differences from rational-mechanical total here expressed in PPQ to milliseconds. The transformed differences are shown in Table 9.

Table 9. Differences from Rational-Mechanical Total Duration.*

<u>Registr.</u>	<u>Difference (PPQ)</u>	<u>Difference (msec)</u>
CV1	-80	-215
CV2	-32	-86
CV3	-9	-24

* Tempo = 93. One PPQ = 2.688172 msec at tempo = 93. The minus sign indicates the actual duration was short of the rational-mechanical norm.

As with the Tape and Pencil method, it is important to keep these differences in perspective. 80 PPQ is short of the duration of a dotted sixteenth note (90), 32 PPQ is just long of a thirty-second note (30), and 9 PPQ is short of a sixty-fourth note (15). Thus, there is not much of an absolute tempo increase in any of the CV's, rather, the slow starts are made up in the long run by a gradual tempo increase, and the net effect amounts to only a fraction of a measure difference from the rational-mechanical total duration.

Results of Studio Session One

There were some technical complications in the first studio session, as related above. Nonetheless, a good amount of usable data was collected. Five versions each for three different drum patterns, and two guitar accompanied performances were registered. There were some false triggered notes in the performances that were unmistakably identified and easily deleted. There were also some missed notes that had to be worked around. Patterns 1 through 3, shown below, were registered at the first session.

The most significant glitch in the data collection for session one was the loss of a good many high-hat notes in the five versions of Pattern 1. Eighth notes on the back beat were lost, but all those on the beat were retained. This was due to the dynamic range problem described in the above section on method: the drummer performed the pattern in such a way that the back beat high-hat hits were in many cases too soft to open the high-hat microphone's gate and register on the sequencer. While this reduces the usefulness of the high-hat registrations for Pattern 1 (a usable number of the hits on the beat were registered), it does not affect the viability of the drum parts. This high-hat problem was corrected for Pattern 2 and Pattern 3 by adjusting the Kepexes and MIDI KITTY.

These complications aside, the data shed more light on how human performances differ from machine performances of rock drum patterns. Following the procedures established in the tape and pencil and pretest portions of the study, the MIDI data for the three patterns were analyzed using SPSS/PC+. Results for each beat will be presented separately.

Before presenting the analysis, a word here about the compilation of the data. For inter-version analyses, the PPQ values of the note durations had to be adjusted because the drummer did not play all the versions at the same tempo. This was accomplished by dividing the average quarter-note duration of a particular version by 240, the sequencer's beat resolution, and then dividing each note duration by this constant. For example, the average quarter-note duration of the high-hat registration for Pattern 2, Version 4 is 228.3. $228.3/240 = .95125$. The first high-hat hit in that version has a duration of 116; therefore, its adjusted duration is $116/.95125$, or, 121.9. The sequencer's tempo setting was 120 when the registration was captured. But the tempo of the performance was 126. In effect, the adjustment is like raising the sequencer's tempo setting to 126 and justifying the note durations to the 240 PPQ beat resolution. The real time duration of the note is not changed significantly. For instance, at tempo = 120, 1 PPQ

= 2.08 msec. $116 * 2.08 = 241.28$ msec = the duration in real time of the high-hat hit. At tempo = 126, 1 PPQ = 1.98 msec. $121.9 * 1.98 = 241.36$ msec. If this procedure were not followed, inter-version analyses, such as a t-test of note durations, would likely be biased by the tempo variations between versions. It also makes possible analysis across the different beats the drummers performed.

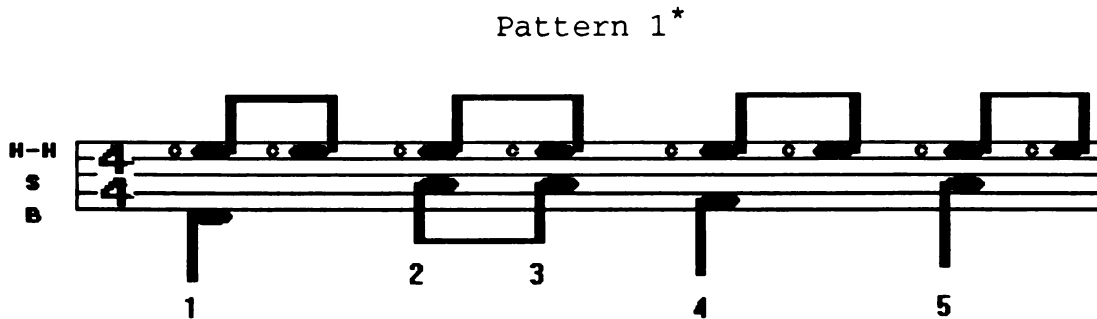


Figure 13. Notation for Pattern 1

*The "c" indicates a closed high-hat sound.

The first step in analyzing the registrations was to check for any systematic variance from the rational-mechanical norm at the note level. As above, this was done with a t-test. For Pattern 1, these procedures were run only on the drum data, excluding the high-hat for the reasons given above. The five versions were each 15 measures long (excluding the count). Consecutive notes with equal rational-mechanical durations were tested. Results are given below.

Table 10. Pattern 1: T-Test of Notes 2 and 3.

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
2	121.67	
3	118.30	.000

Table 11. Pattern 1: T-Test of Notes 4 and 5.*

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
4	239.67	
5	241.95	.000

* This test was run excluding measure eight, because the drummer always performed a fill over note five of measure eight.

Table 12. Pattern 1: T-Test of Notes 5 and 1.*

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
5	241.95	
1	239.09	.000

* This test was run excluding measure eight, because the drummer always performed a fill over note five of measure eight.

Tables 10-12 indicate that, on average, consecutive notes with equal rational-mechanical durations were performed with unequal actual durations, and that these differences are probably not due to chance.

Of particular interest is the difference for notes 5 and 1. The long-short relationship extends across the measure boundary. Since the accents fall on beats one and four in this pattern, the difference might indicate a durational accent on beat one. Beat three, however, does not appear to have a durational accent, because the mean durations for beats two and three are 239.97 and 239.67, respectively, both essentially equal to the rational-mechanical value of 240.

This is not to say, however, that beat three is not accented. The velocity values for notes 1 and 4 (beats one and three) show a pattern of stress accents, where on average beat three is hit harder than beat one. Table 13 gives the results of a t-test on the mean velocity values for beats one and three.

Table 13. Pattern 1: T-Test of Velocity for Notes 1 and 4.

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
1	75.41	
4	80.57	.000

The t-test results show the way actual durations were performed on average. However, it is important to keep in mind the actual durations of, say, notes 2 and 3 were not

always performed in a long-short fashion. Furthermore, while there are very few instances where consecutive notes of equal rational-mechanical duration have equal actual durations, often times the actual differences are quite small. For Pattern 1, there are forty-two instances over the fifteen measures where consecutive notes of equal rational-mechanical durations occur. By counting the number of cases where the PPQ difference would equal or exceed 10 msec, an index can be computed that indicates how often in a given version variance between consecutive notes of equal rational-mechanical duration is likely to be of consequence to the listener. Table 14 gives this computation for each of the five versions.

Table 14. Pattern 1: Number of Consequential Differences.

<u>Version</u>	<u>Cases (out of 42)</u>	<u>% Consequential*</u>
1	14	33
2	18	43
3	18	43
4	15	36
5	16	38
	-----	-----
MEAN	16	38

* Based on the 10 msec threshold for human perception of durational differences given by Seashore.

Table 14 shows that, of the forty-two possible places where systematic variance could be applied to consecutive notes of equal duration, variance of a magnitude likely to be of consequence to the listener was applied only between thirty-three and forty-three percent of the time. This works out to an average of thirty-eight percent. Another way of looking at this is that sixty-two percent of such differences in Versions 1 through 5, though actually different, can be considered virtually equal.

Looking at the five versions of Pattern 1 in terms of macro tempo change, there are some interesting findings. Tables 15 and 16 give the regression line slopes and intercepts and the average tempos for the five versions.

Table 15. Pattern 1: Slopes and Intercepts.*

<u>Vers.</u>	<u>Slope</u>	<u>Intercept</u>
1	1.00155	-19.85
2	.99804	-5.05
3	.99264	.39
4	.99985	-21.96
5	1.00195	-16.24

* Computed at the measure level from adjusted progress scores.

Table 16. Average Tempos for Pattern 1, Versions One Through Five.

<u>Version</u>	<u>Avg. Tempo</u>
1	125
2	124
3	125
4	124
5	124

Table 15 shows that two of the versions (1 and 5) have tempo decreases at the macro level, while three of them (2, 3 and 4) have tempo increases. Versions 1 and 5 start fast and gradually slow down, and version 4 starts fast and gradually speeds up.

The average tempo values are interesting because of the fact that the drummer was given a four beat count at tempo = 120 just before performing each of the versions. Though he was given this count, he chose to play the pattern at a tempo of 124 or 125. This says something about how accurately Drummer A was able to perceive the tempo from four beats. It may also mean that the arbitrarily set tempo of 120 felt too slow for the drummer for this particular pattern.

A last way of looking at Pattern 1 is in terms of the residuals of the regression equation. The residuals are the differences between the scores predicted by the regression model and the actual durations performed by the drummer. In some cases, the regression models were surprisingly accurate

in predicting the actual progress scores. In other cases, there were sizable differences. Upon examining plots of the residuals, a pattern was noticed to be more or less consistent between them. This similarity raised the question of whether a non-linear model might fit better to the data.

Figure 15 (next page) shows the residual scores for each version plotted against the number of measures. Each plot has the form of an inverted hump. Using the Trends Module in SPSS/PC+, it was possible to fit different non-linear models to the data, in an attempt to find an equation which generally represents the tempo change over the course of 15 measures.

A quadratic model made a great improvement in the accuracy of predicting the actual progress scores. Table 17 shows how this improved accuracy for Version 1.

Table 17. Pattern 1, Version 1: Comparison of the Accuracy of Linear and Quadratic Models in Macro tempo Prediction.

<u>Measure</u>	<u>Linear Model Error</u>	<u>Quadratic Model Error*</u>
1	17	2
2	7	-2
3	-2	-5
4	1	2
5	-4	1
6	-3	4
7	-7	2
8	-13	-4
9	-12	-3
10	-7	0
11	-8	-3
12	-1	0
13	5	2
14	12	3
15	13	-3
	-----	-----
TOTAL	112	35

* Error is measured in PPQ. A positive error value means the predicted duration was short of the actual duration; a negative error value means the opposite.

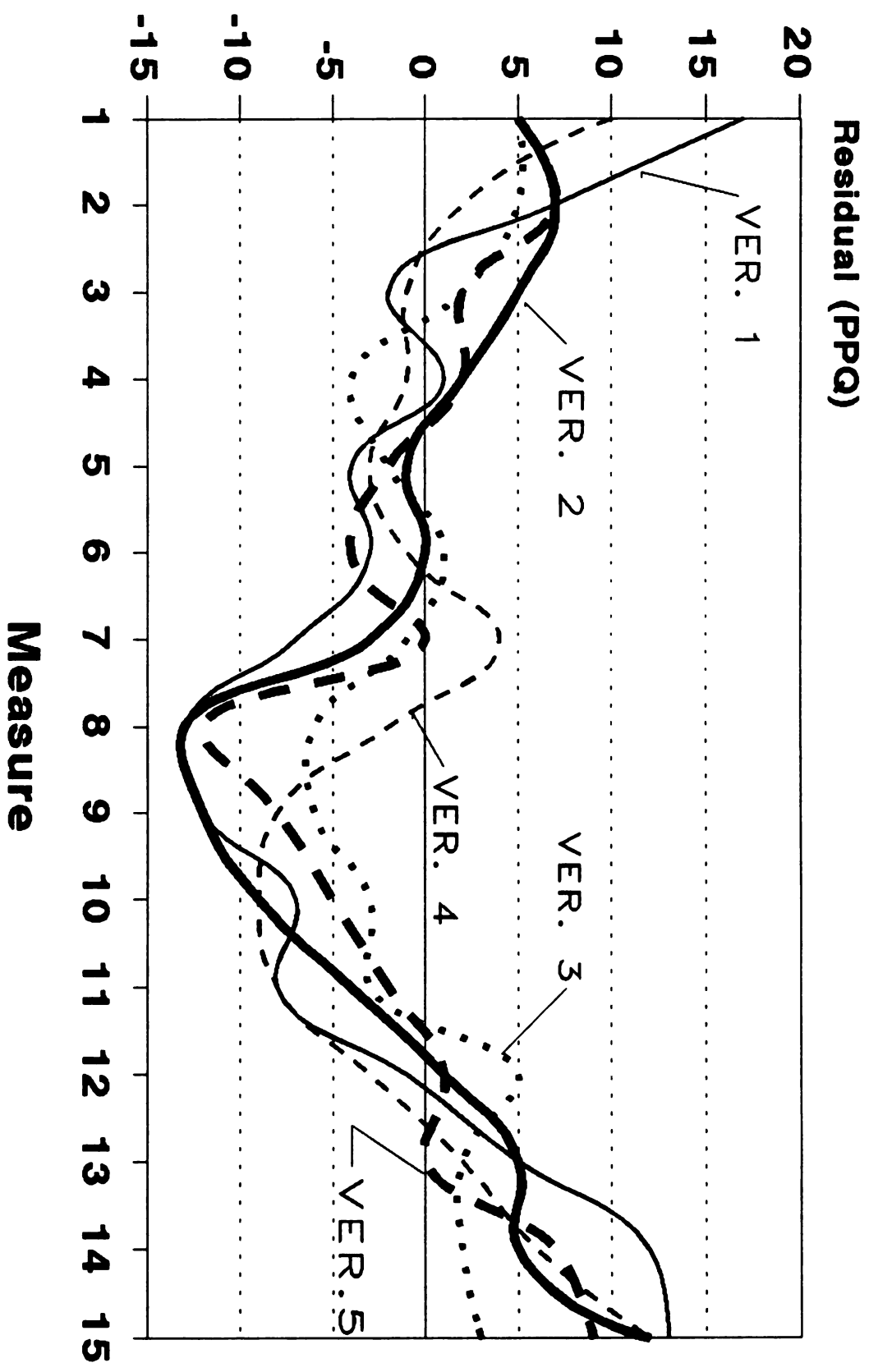


Figure 15. Residual Maps for Pattern 1

Table 17 illustrates the improved accuracy of the quadratic model in describing the macro tempo change in Version 1. In this case, a sixty-nine percent reduction in total error was achieved by use of a non-linear model. Macro tempo in Version 1 is characterized by a fast start and an overall slowdown. The non-linear model is more accurate because the slow down in Version 1 comes at the end. Macro tempo is running ahead of the rational-mechanical norm through measure twelve. The overall tempo decrease, represented by a total duration greater than the rational-mechanical norm, is accomplished in the last three measures. The quadratic model better fits this precipitous drop in tempo. Thus, the drastic reduction in error. A similar reduction of fifty-two percent was found for Version 5, which in macro tempo is very similar to Version 1. For Versions 2-4, the improvements in accuracy due to use of the quadratic model ranged from thirty-nine to forty-six percent. It should also be observed that in every case the quadratic model significantly narrowed the range of the error. This means that individual errors are smaller for the quadratic model, as well as, total error, which is further support for its improved accuracy.

Pattern 2*

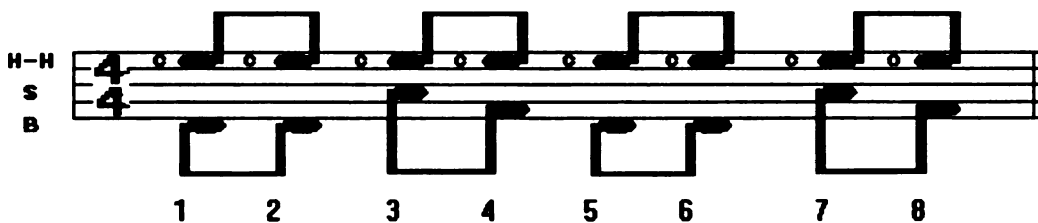


Figure 16. Notation for Pattern 2

*The "c" indicates a closed high-hat sound, "o" indicates open high-hat sound. In performing the pattern, Drummer A used only a closed high-hat sound, the transcription represents the pattern as taken from the drum machine.

Pattern 2 is made up entirely of eighth-notes, and was chosen for this fact. As with the other beats, the drummer was instructed to place a fill at measure eight, beat four. Beside this, the five versions of Pattern 2 were made up exclusively of eighth-notes.

T-tests were performed on the mean values for consecutive notes of equal rational-mechanical duration. The results are given in Table 18 below.

Table 18. Pattern 2: T-Tests of Consecutive Notes with Equal Rational-Mechanical Durations.*

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
1	119.63	
2	118.66	.026+
3	121.16	.000+
4	120.49	.104
5	119.87	.137
6	118.27	.000+
7	121.42	.000+
8	120.49	.031+
1	119.63	.016+

* Computations for notes 7 and 8 exclude measure eight because of fill.

+ Probability of .05 level or better.

On average, Pattern 2 seems to have been performed as a series of two note groupings, each having a long-short duration orientation. The differences were large enough to reject the null hypothesis that they are due to chance in every case except for Notes 3 and 4.

Pattern 2 was also examined at the quarter-note level, to see if durational accents were prevalent. Evidence was found of a lengthening, on average, of beats two and four and a shortening of beats one and three, similar to a pattern found for beats four and one in the above analysis of Pattern 1. The results are given below in Table 19.

Table 19. Pattern 2: T-Tests of Beat Durations, Drums.

<u>Beat Number</u>	<u>Mean</u>	<u>Significance</u>
1	238.66	
2	241.57	.000
3	238.69	.000
4	241.25	.000
1	238.66	.000

Adjustments to the Keypexes and MIDI KITTY made the high-hat note registrations for Pattern 2 consistent enough to permit an analysis. Unfortunately, these adjustments reduced the dynamic range so far as to sacrifice the velocity data: there is not enough variance in the velocity data for Pattern 2 and Pattern 3 to analyze. Thus, a sort of trade off was made between usable velocity data and unusable high-hat duration data in Pattern 1 and the reverse in Pattern 2.

The high-hat durations were analyzed using much the same methods as were used for the drum data. First, t-tests were run on the eight notes making up the high-hat pattern. The results are given in Table 20:

Table 20. Pattern 2: T-Tests on High-Hat Note Durations.

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
1	120.93	
2	118.24	.000
3	120.68	.000
4	120.85	.695
5	120.73	.776
6	117.78	.000
7	121.17	.000
8	120.30	.065
1	120.93	.143

Compared to Table 18, the high-hat duration t-tests appear quite similar. Once again, there is no significant difference between the means of Notes 3 and 4. However, Notes 1 and 8 were played longer on average for the high-hat. As a result, the differences between Notes 7 and 8, and Notes 8 and 1 cannot be said to be significant.

Though they differ somewhat, the mean duration values for the high-hat and drums are similar enough that one would think them to have a fairly high correlation. Their averages agree in terms of direction, if not magnitude, in six of the eight points where the t-test checked the differences in duration. However, as Table 21 shows, the correlation are not high.

Table 21. Pattern 2: Correlations Between Drum and High-Hat Durations.*

<u>Version</u>	<u>Pearson Correlation Coefficient</u>
1	.31
2	.24
3	.24
4	.22
5	.54
All	.32

* Versions 1-5 were computed using the unadjusted PPQ values, since these are intra-version computations. "All" was computed using the adjusted scores because it is an inter-version computation. All of the coefficients were significant to at least the .01 level.

A look at the raw data from Pattern 2's registrations reveals many instances when, for a pair of eighth notes, the drum performance is long-short and the high-hat is short-long, or vice-versa. This sort of occurrence can lower the correlation coefficient, which is sensitive to both the magnitude and the direction of variation from the mean drum and high-hat durations. In fact, the drum to high-hat correlation and the mean duration for a note in a given pattern can go in opposite directions.

A short example points this out. Consider the first two measures of Pattern 2, Version 4:

Table 22. Durations for the First Two Measures of Pattern 2

<u>Measure</u>	<u>No. of Note</u>	<u>Drum Dur.</u>	<u>High-Hat Dur.</u>
1	1	116	116
1	2	117	116
1	3	116	115
1	4	113	115
1	5	114	114
1	6	111	109
1	7	113	115
1	8	116	114
2	1	112	117
2	2	115	112
2	3	114	119
2	4	117	113
2	5	116	115
2	6	110	108*
2	7	115	118
2	8	116	113

As is, the drum to high-hat correlation for these two measures is .43 and the mean duration for high-hat note 6 in the pattern is 109. If the duration of high-hat note 6 in measure 2 (marked with "**") is changed to 118, making the relationship between high-hat Notes 5 and 6 in measure 2 short-long instead of long-short, the correlation drops to .06, though the mean duration for high-hat note 6 increases to 114.

Thus, in describing the inter-relation between the drum and high-hat notes for Pattern 2, and in other cases where duplicate rhythm patterns are being performed simultaneously, the correlation statistic is not very useful and perhaps even misleading. A simpler and better approach is to look at a frequency distribution of the offset between notes that are to occur simultaneously, according to the rational-mechanical norm. From this view, the durations of the high-hat notes are less important than when their start times occur relative to the start times of the drum hits. Table 23 shows the offsets for Version 1 of Pattern 2:

Table 23. Pattern 2, Version 1: Drum/High-Hat Offsets.*

<u>Value</u>	<u>Frequency</u>	<u>Percent</u>	<u>Cum. Percent</u>
-6	2	1.6	1.6
-5	2	1.6	3.3
-4	7	5.7	8.9
-3	5	4.1	13.0
-2	18	14.6	27.6
-1	15	12.2	39.8
0	38	30.9	70.7
1	7	5.7	76.4
2	14	11.4	87.8
3	7	5.7	93.5
4	5	4.1	97.6
5	1	.8	98.4
6	1	.8	99.2
9	1	.8	100.0

TOTAL	123	100.0	

* Values are in unadjusted PPQ units. One PPQ = 2.08 msec. Therefore, a value of "0" means that the drum and high-hat start times were closer than 2.08 msec, or, less than 1 PPQ. A negative value means that the drum hit before the high-hat; a positive value means the reverse.

From Table 23 it is learned that most of the time (71%) the drum hit before the high-hat or less than 2.08 msec apart from it. The range of offset is from -6 to +9 PPQ, with a standard deviation of 2.4. This means that, generally, the offset will fall in the +/- 2 PPQ range, according to the pattern established in Pattern 2, Version 1.

A 2 PPQ offset amounts to about 4 msec. Listeners may not be able to discern the delay; however, they may perceive the aural effect of offsetting the drum and high-hat notes by 4 msec or so. The drummer may be doing something similar to chord serialization, where the attacks of notes in a chord are spread out in time by the performer to display its structure or bring out the melody. Likewise, the drummer may be trying to bring out the sound of the drum or the high-hat by making one occur slightly before the other. Consequently, offsets of less than 10 msec should not be viewed in the same way as durational differences of that magnitude in two notes of a series. This is because it is probably easier for a listener to perceive a small offset than an equally small durational difference. Of course, such an assertion should be checked out. A test of this will be included in the audience test.

While the results from Version 1 are interesting, a more general view of the offsetting in Pattern 2 is had by looking at a frequency distribution incorporating all five

versions. It is also possible to look at sub-distributions showing offsetting for the bass drum and snare drum individually. Tables 24-28 give these distributions and sub-distributions and their summary statistics.

Table 24. Pattern 2: Drum/High-hat Offsets, Bass and Snare Drum.*

<u>Value</u>	<u>Frequency</u>	<u>Percent</u>	<u>Cum. Percent</u>
-7	5	.8	.8
-6	8	1.3	2.2
-5	14	2.3	4.5
-4	33	5.5	10.0
-3	37	6.2	16.2
-2	115	19.2	35.3
-1	115	19.2	54.5
0	126	21.0	75.5
1	40	6.7	82.2
2	39	6.5	88.7
3	35	5.8	94.5
4	26	4.3	98.8
5	1	.2	99.0
6	3	.5	99.5
7	1	.2	99.7
9	1	.2	99.8
12	1	.2	100.0

TOTAL	600	100.0	

* Based on unadjusted PPQ values.

Table 25. Pattern 2: Drum/High-hat Offsets, Bass Drum Only.*

<u>Value</u>	<u>Frequency</u>	<u>Percent</u>	<u>Cum. Percent</u>
-7	5	1.1	1.1
-6	8	1.8	2.9
-5	14	3.1	5.9
-4	31	6.8	12.8
-3	36	7.9	20.7
-2	87	19.2	39.9
-1	66	14.5	54.4
0	66	14.5	68.9
1	37	8.1	77.1
2	36	7.9	85.0
3	35	7.7	92.7
4	26	5.7	98.5
5	1	.2	98.7
6	3	.7	99.3
7	1	.2	99.6
9	1	.2	99.8
12	1	.2	100.0

TOTAL	454	100.0	

* Based on unadjusted PPQ values.

Table 26. Pattern 2: Drum/High-hat Offsets, Snare Drum Only.*

<u>Value</u>	<u>Frequency</u>	<u>Percent</u>	<u>Cum. Percent</u>
-4	2	1.4	1.4
-3	1	.7	2.1
-2	28	19.2	21.2
-1	49	33.6	54.8
0	60	41.1	95.9
1	3	2.1	97.9
2	3	2.1	100.0

TOTAL	146	100.0	

* Based on unadjusted PPQ values.

Table 27. Pattern 2: Offset Summary Statistics For Overall, Bass Drum, and Snare Drum Distributions.

<u>Dist.</u>	<u>Range</u>	<u>Mode</u>	<u>Standard Deviation</u>
Overall	19	0	2.394
Bass Dr.	19	-2	2.696
Snare Dr.	6	0	.978

Table 28. Pattern 2: Proportion of Time Drum Hits Before, After, or Simultaneous with High-Hat.

<u>Dist.</u>	<u>% Simultaneous*</u>	<u>% Before H-H</u>	<u>% After H-H</u>
Overall	21	54.5	24.5
Bass Dr.	14.5	54.4	31.1
Snare Dr.	41.1	54.8	4.1

* "Simultaneous" means that the drum and high-hat hits were within 2.08 msec of each other.

These tables reveal several interesting facts. First, most offsets for the bass tend to fall in the +/- 3 PPQ range, while for the snare drum the range is only +/- 1 PPQ. The most frequent offset for the overall and snare drum distributions is 0, and for the bass drum it is -2. The most frequent offset (mode) for the snare drum makes up 41.1% of its distribution. For the bass drum, the mode accounts for only 14.5% of the distribution, pointing out that the bass drum offsets have greater variability than for the snare. Most of the time, the drummer did not hit the drum and high-hat simultaneously. In the majority of cases, the drummer hit the drum before or at the same time as with the high-hat. For the snare drum, however, there are relatively few instances where the high-hat is hit greatly before the drum, his preference was to hit them simultaneously or to hit the snare only slightly earlier.

As for Pattern 1, the number of duration differences likely to be consequential to the listener were counted. In general, there were lower percentages of consequential differences than for Pattern 1. Table 29 gives these results.

Table 29. Pattern 2: Number of Consequential Differences.

<u>Version</u>	<u>% Conseq. (Drums)</u>	<u>% Conseq. (H-H)</u>
1	25	29
2	12	25
3	25	27
4	20	22
5	30	30
	-----	-----
Mean	22	27

* Duration differences are considered consequential (as in % Conseq.) if they equal or exceed 10 msec. Percentages are out of 118 possible instances.

The macro tempo changes for the five versions of Pattern 2 were computed using the established procedure from the adjusted progress scores. The slopes and intercepts are given in Table 30 and Table 31.

Table 30. Pattern 2: Slopes and Intercepts*

<u>Version</u>	<u>Slope</u>	<u>Intercept</u>
1	.99667	18.067
2	.99868	5.409
3	.99697	24.923
4	1.00193	5.409
5	.99940	31.734

* Computed at the measure level from adjusted high-hat progress scores.

In terms of macro tempo change, Pattern 2 contrasts from Pattern 1 in that the intercepts are positive. Versions 2 and 4 are within six PPQ of a zero intercept, meaning that they begin at a rate near the rational-mechanical tempo. Versions 1, 3 and 5, however, have rather large positive intercepts, indicating a considerable slow start. All versions have gradual tempo increases, except Version 4. Version 4 starts near the rational-mechanical tempo and gradually slows down. The slow starts are made up for by the increases in all but Version 5, which, in spite of its tempo increase, has a total duration longer than the rational-mechanical norm.

The average tempos for the five versions are given in Table 31. Again, the drummer did not follow exactly the tempo = 120 count he was given, rather, he chose to perform the pattern at higher tempos.

Table 31. Pattern 2: Average Tempos for the Five Versions.

<u>Version</u>	<u>Average Tempo</u>
1	128
2	127
3	126
4	126
5	126

A last way of looking at Pattern 2 is in terms of its residual maps. The residual maps for Pattern 2 are much flatter than for Pattern 1. They are also inverted relative to Pattern 1. In general, these residual maps hold closely to the zero line, meaning that the macro tempo increases or decreases in Pattern 2 are fairly linear. Therefore, no attempt was made to fit a non-linear model to the data for Pattern 2. Figure 17 (next page) shows the residual maps for Pattern 2.

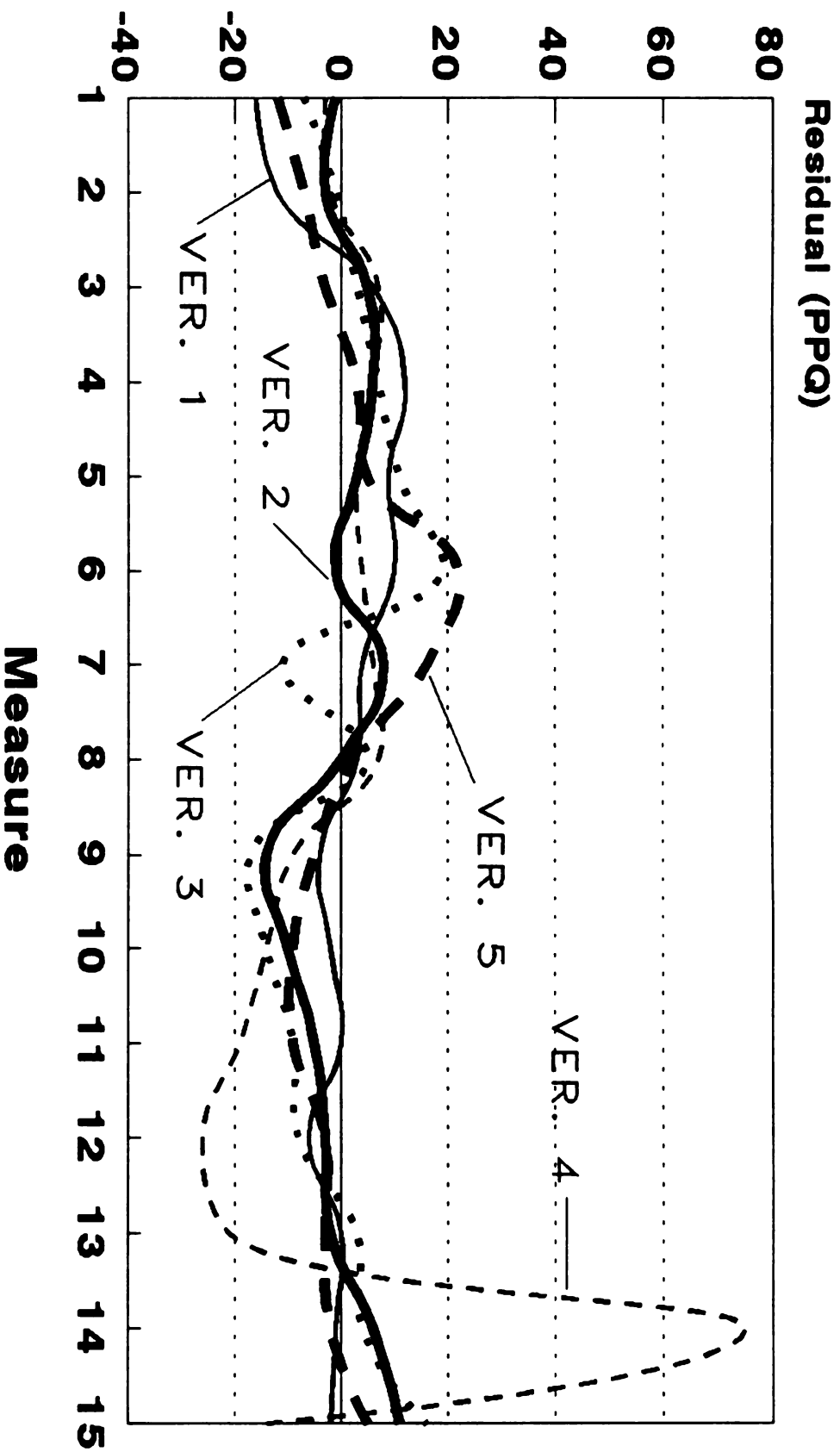
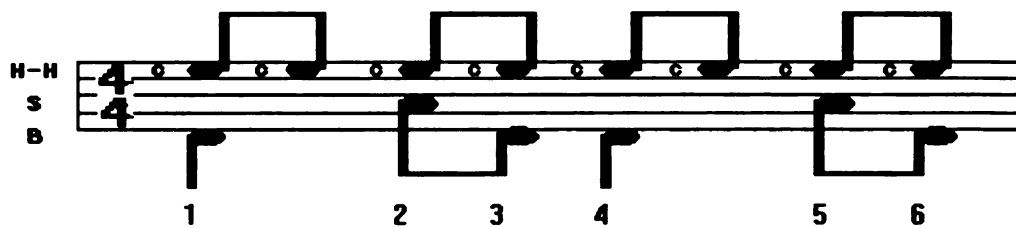


Figure 17. Residual Maps for Pattern 2

Pattern 3

Measure A



Measure B

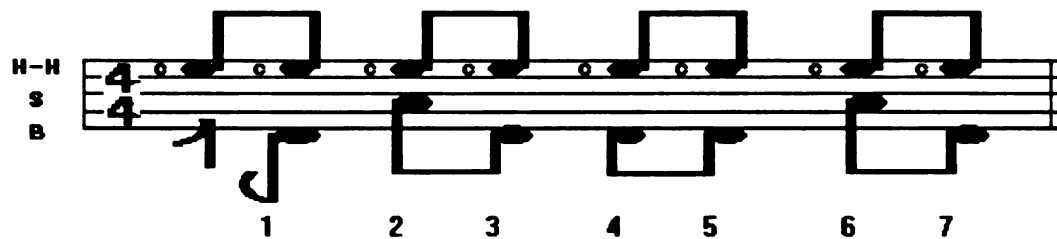


Figure 18. Notation for Pattern 3

Pattern 3 is distinct from Pattern 1 and Pattern 2 in that its pattern extends over two measures, instead of one. It is further distinguished by containing a rest. A point of similarity, though, is the constant eighth note figure in the high-hat, allowing some comparison with the high-hat performances of Pattern 2.

The first procedure run on Pattern 3 was a battery of t-tests, as usual, to look for systematic variance in consecutive notes of equal rational-mechanical duration. The difference between the means for notes 2 and 3 was first checked. This snare/bass drum eighth note figure is also found on the second beat of the pattern in Pattern 2. Here, unlike in Pattern 2, it was performed long-short, and the difference between means was significant. Table 32 gives the results of this t-test.

Table 32. Pattern 3: T-test of Notes 2 and 3.*

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
2	122.20	
3	121.18	.005

* based on adjusted duration values. n=74.

The above t-test was made across both measures of the two measure pattern making up Pattern 3. Looking at only measure B, it is possible to test for significant differences between the means of other pairs of eighth-notes. Table 33 gives the results of these tests.

Table 33. Pattern 3, Measure B: Consecutive Note T-Tests*

<u>Note Number</u>	<u>Mean</u>	<u>Significance</u>
3	120.24	
4	119.67	.298
5	116.93	.000
6	121.83	.000
7	121.06	.163

* Based on adjusted duration values, computed excluding measure 8 because of fill. n=30.

Table 33 reveals a long-short-long orientation between notes 4, 5 and 6 that contains significant differences in mean duration. The pair of eighth note bass drum hits, notes 4 and 5, are performed long-short on average.

An interesting feature of Pattern 3 is the rest at the beginning of measure B. It displaces the normal down-beat accent on beat one to the back-beat. The rational-mechanical distance between the start time of note 6 in measure A and note 1 in measure B is equal to 240, or, one quarter-note. A rough way to gauge how the drummer is performing the rest can be had by comparing this start time distance to the duration of note 1 in measure A. Table 34 gives the result of a t-test between the duration and the start time distance.

Table 34. Pattern 3: T-Test to Gauge the Performance of the Rest in Measure B.*

<u>Duration</u>	<u>Mean</u>	<u>Significance</u>
Note 1A	237.57	
Note 6A-1B	244.70	.000

* Based on adjusted duration values. n=40.

The durational difference shown in Table 34 amounts to a sizable delay preceding Note 1 of Measure B. It was theorized above that delaying a note is a sort of durational accent. It is also possible that drummers treat rest durations differently than note durations, regardless of whether they displace normal down-beat accents.

The high-hat figures in Pattern 2 and Pattern 3 are musically identical in a notational sense. This begs the question of whether identical high-hat figures in different drum beats are performed in a similar way. They would be performed exactly the same according to the rational-mechanical norm. To test the degree of similarity in the high-hat performances of Pattern 2 and Pattern 3, the correlation statistic was used. The correlation coefficient (r) between the adjusted high-hat durations of Pattern 2 and Pattern 3 was .08, a very weak correlation.

A benchmark for the weakness of this correlation can be had by considering the correlation between the drums and high-hat for Pattern 2. Recall that Pattern 2 was made up of all eighth-notes, in both the drum and high-hat parts. The correlation between the drum and high-hat durations across the five versions of Pattern 2 was .32. The disparity between .08 and .32 suggests interaction between the high-hat and drum parts. Though notationally identical, the high-hat performances between Pattern 2 and Pattern 3 are rendered quite differently in terms of actual durations.

Another way to compare Pattern 2 and Pattern 3 is in terms of offsets between drum and high-hat start times. Tables 35 and 36 give the summary statistics for the overall, bass and snare drum offset distributions for Pattern 3.

Table 35. Pattern 3: Offset Summary Statistics For Overall, Bass Drum and Snare Drum Distributions.

<u>Dist.</u>	<u>Range</u>	<u>Mode</u>	<u>Standard Deviation</u>
Overall	20	0	2.607
Bass Dr.	20	3	2.864
Snare Dr.	6	-2	.973

Table 36. Pattern 3: Proportion of Time Drum Hits Before, After, or Simultaneous with High-Hat.*

<u>Dist.</u>	<u>% Simultaneous*</u>	<u>% Before H-H</u>	<u>% After H-H</u>
Overall	19.3	40.5	40.1
Bass Dr.	14.1	29.0	56.9
Snare Dr.	31.3	66.7	2.0

* "Simultaneous" means that the drum and high-hat hits were within 2.08 msec of each other.

The offset summary statistics for Pattern 3 have some interesting similarities and differences to those of Pattern 2. First, they are very close in terms of range. The overall range of 20 for Pattern 3 compares to 19 for Pattern 2, the bass drum range of 20 compares to 19, and the snare drum ranges are equal at 6. The standard deviations are also quite similar. An interesting difference is that in Pattern 3, for the bass drum, the drummer favored hitting the bass drum after the high-hat over half the time. This is a reversal from the pattern in Pattern 2. However, the Pattern 3 snare drum offsets look very similar to the pattern in Pattern 2: the majority of the cases have the snare hitting slightly before the high-hat, and there is again a relatively greater incidence of simultaneous hits.

Comparison in terms of duration differences equal to or exceeding 10 msec shows a similarity between Pattern 3 and Pattern 2. Table 37 gives the percentages by version for the drum and high-hat portions of the registrations.

Table 37. Pattern 3: Number of Consequential Differences.*

<u>Version</u>	<u>% Conseq. (drums)</u>	<u>% Conseq. (high-hat)</u>
1	14	37
2	30	27
3	34	35
4	24	37
5	36	30
	-----	-----
Mean	28	33

*Drum figures out of 50 possible cases, high-hat out of 118.

The mean number of duration differences equal to or greater than 10 msec of twenty-eight and thirty-three percent for drums and high-hat, respectively, in Pattern 3 compares with twenty-two percent for drums and twenty-seven percent for high-hat in Pattern 2. Though the averages for Pattern 3 are higher overall, they agree in magnitude and direction with Pattern 2. The Pattern 1 average for drums of thirty-eight percent is greater than for Pattern 2 or 3.

Average tempos for the five versions of Pattern 3 were again all above the given count of 120. Table 37 lists these average tempos.

Table 38. Pattern 3: Average Tempos.

<u>Version</u>	<u>Tempo</u>
1	126
2	125
3	124
4	124
5	124

Table 39. Pattern 3: Slopes and Intercepts

<u>Version</u>	<u>Slope</u>	<u>Intercept</u>
1	1.00123	-18.96
2	.99905	-18.09
3	1.00165	-28.45
4	.99947	.79
5	1.00074	.89

Table 39 indicates that versions 1-3 had fast starts, while Versions 4 and 5 started very near the rational-mechanical origin of 0. Three of the versions (1, 3 and 5) have gradual tempo decreases, two of them (2 and 4) have increases.

Figure 19 (next page) shows the residual maps for Pattern 3. The residuals range from 0 to 28 PPQ. A thirty-second note equals 30 PPQ. Therefore, the worst case error in predicting the actual progress scores amongst the five equations is less than a thirty-second note.

The shape of the residual maps for Versions 1 and 3 appear similar to those of Pattern 1. Also, the wider range of the residuals in these two versions makes one wonder whether a non-linear equation might better describe their macro tempo changes. The quadratic equations produced significant improvements in predicting the actual progress scores for those versions. The improvements were a sixty-two percent reduction in total error for Version 1 and a sixty-three percent reduction for Version 3. These improvements were also attended by greatly reduced error ranges.

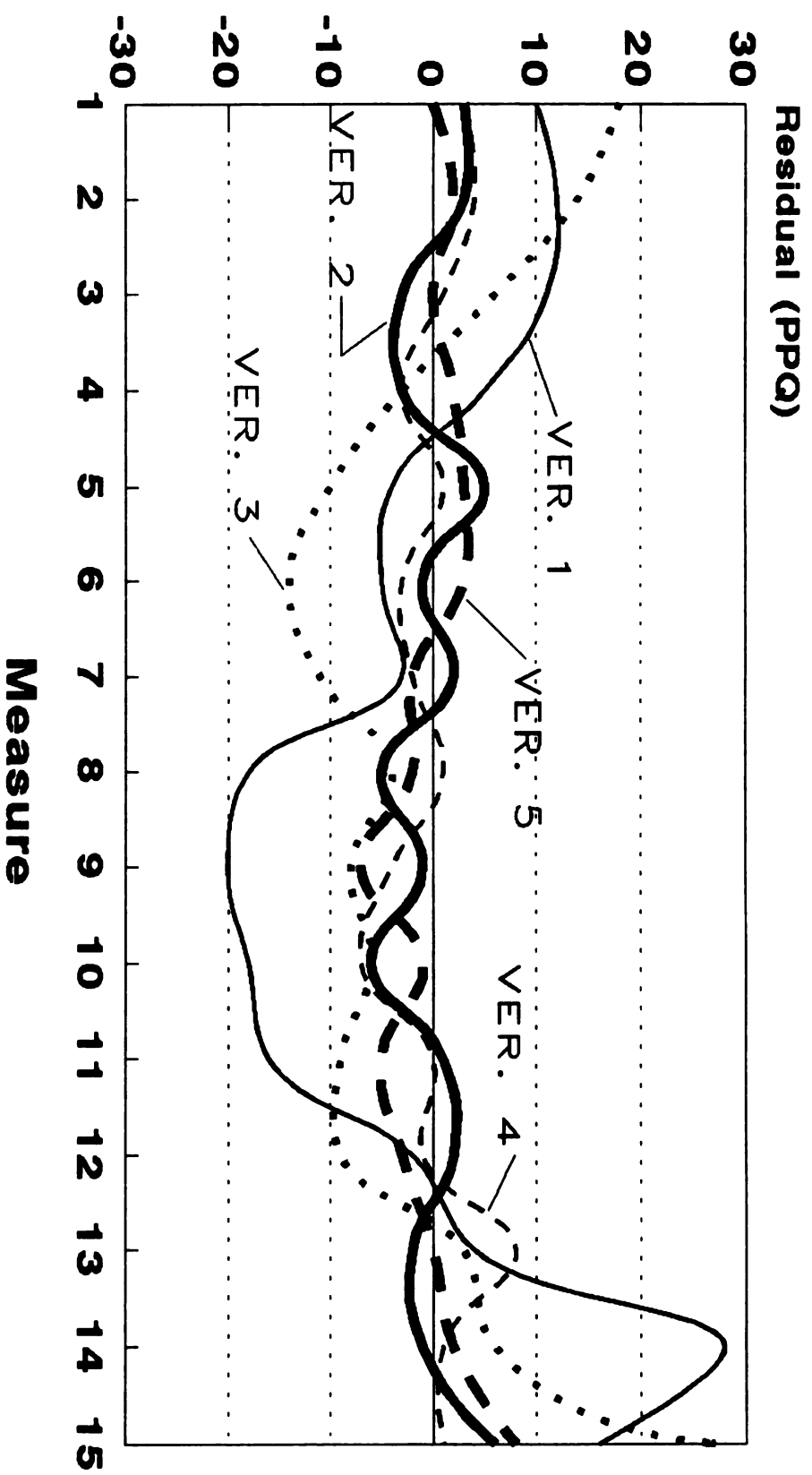


Figure 19. Residual Maps for Pattern 3

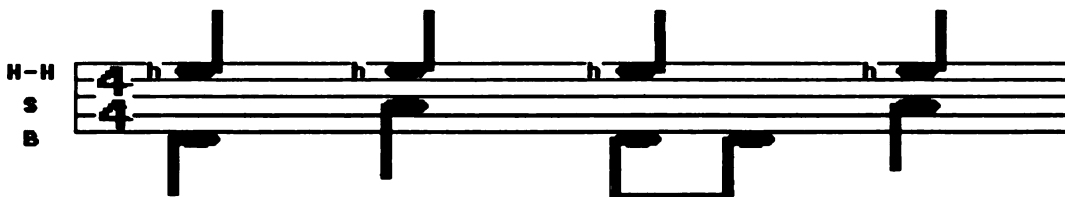
Results of Studio Session Two

The goal of the second studio session was primarily to obtain performances for use in the audience test portion of the study. The second drummer was not asked to do as many repetitions of each pattern he performed. Thus, there are certain statistical procedures, such as t-test or factor analysis, that cannot be performed on the data from session two, because the relatively small sample size would make even large differences in average note durations appear insignificant. Also, because of the way Drummer B set up and played his equipment, no velocity data are available for session two. Despite this, there is still an opportunity for a limited analysis of Drummer B's performances in order to gain more knowledge about rock drumming, as well as, to compare him to Drummer A.

As with the analysis of the first session, the different patterns performed by Drummer B will be dealt with individually in the following.

Pattern 4*

Measure A



Measure B

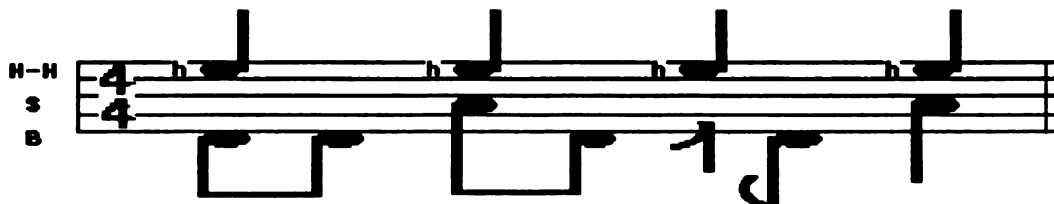


Figure 20. Notation For Pattern 4

*The "h" indicates a half-closed high-hat sound. Drummer B used a full-closed sound in his performances.

Like Pattern 3, Pattern 4 is a two measure pattern. It also contains a rest in Measure B. There were two versions of Pattern 4 performed, but version 1 was unusable because of a computer error in writing the track in the song file. A unique feature of Pattern 4 is the quarter-note figure in the high-hat. It offers an interesting comparison to Patterns 2 and 3 in terms of offsetting. Pattern 4 was recorded with the sequencer's tempo setting at 130. This was done because the pattern sounded unnaturally slow at tempo = 120. Raising the sequencer's tempo setting improves the accuracy of the registration; for Pattern 4 it was +/- 1.92 msec. Tables 40-41 give the summary statistics for offsetting in Pattern 4.

Table 40. Pattern 4: Summary Statistics for the Over-all, Bass and Snare Drum Distributions.

<u>Dist.</u>	<u>Mode</u>	<u>Range</u>	<u>Standard Dev.</u>
Overall	0	12	3.336
Bass Dr.	-7	9	2.694
Snare Dr.	0	8	2.113

Table 41. Pattern 4: Percentage of Simultaneous and Non-Simultaneous Hits by Drum.

<u>Drum</u>	<u>% Simultaneous</u>	<u>% Before</u>	<u>% After</u>
Bass	12.5	83.3	4.2
Snare	19.9	22.6	58.1

There is a good deal of variance in offsetting in Drummer B's rendition of Pattern 4. The mode makes up for only sixteen to twenty percent of the respective distributions. There is also a discernible difference in offsetting between the bass and snare drum. The trend is to hit the bass drum before the high-hat, and the high-hat before the snare. Since the overall distribution encompasses only 55 values, it is possible that the observed trend is merely a chance occurrence. A way to check against this possibility is to use the chi-square statistic. It provides a level of certainty for saying there is a trend in a distribution of only 55 values. According to the chi-square statistic, the probability is very good ($p > .0000$) that there is a systematic difference in offsetting between the two drums in Pattern 4.

The regression equation for Pattern 4 has a slope of .99723 and an intercept of 2.21. Thus, it begins very close to rational-mechanical tempo and speeds up from there. The range of the residuals is 0 to 20 PPQ, 20 PPQ being less than a thirty-second note's duration. The appearance of Pattern 4's residual map, depicted in Figure 21, resembles those of Pattern 1.

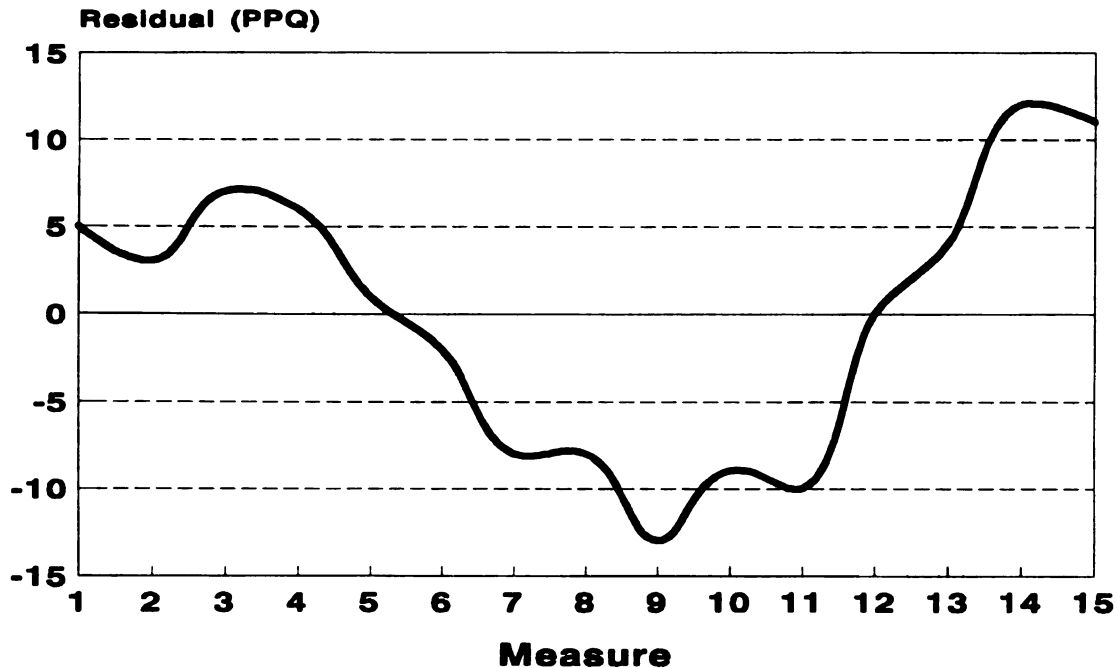


Figure 21. Residual Map for Pattern 4

It is possible to go a step farther than visual comparison of the residual maps of Patterns 1 and 4. Correlating their residual scores gives a numerical indicator of the strength of any relationship. Further, comparing the Pattern 1 and Pattern 4 residual correlations with those of Pattern 4 and other less visually similar residual maps, one gets a relative measure of how closely Pattern 4's residual scores resemble Pattern 1's. Such a comparison is shown in Table 42.

Table 42. Correlations Comparing Pattern 1, 2, and 4 Residual Scores.

<u>Version</u>	<u>Pattern 1 w/ Pattern 4</u>	<u>Pattern 2 w/ Pattern 4</u>
1	.8477**	-.0381
2	.8905**	.6739*
3	.6580*	.5963*
4	.7298*	.4005
5	.8165**	-.0116

* = $p > .01$ ** = $p > .001$

What Table 42 shows is the similarity, in terms of non-linearity, between Pattern 4 and Patterns 1 and 2. The strongest similarities are between Pattern 4 and Pattern 1, versions 1, 2 and 5. The relationship between Pattern 4 and Pattern 2 is weaker. However, it is remarkable that Pattern 2 has two versions (2 and 3) with fairly high correlations to Pattern 4. This suggests that the non-linearity found amongst the various beats and versions may have some coherence, even falling into several different types (perhaps, another type represented by the low correlations), that may be useful in synthesizing rock drum beats.

Pattern 5

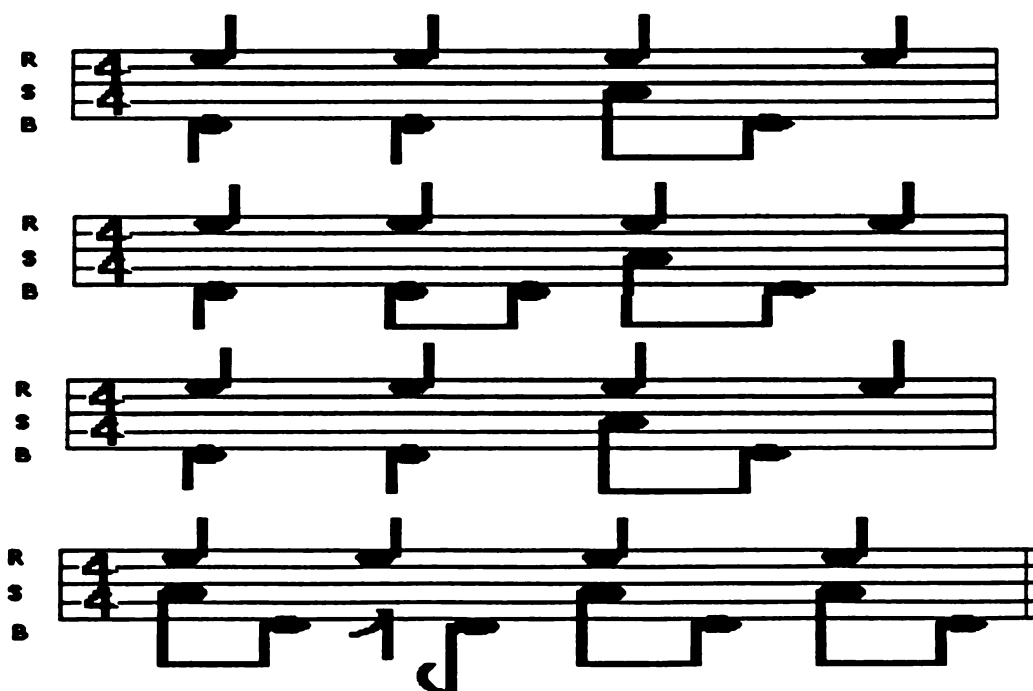


Figure 22. Notation for Pattern 5

Pattern 5 is a four measure pattern. It contains a rest and employs a ride cymbal (R) instead of a high-hat. Pattern 5 was recorded with the sequencer's tempo setting at 144. Again, this was done because the pattern seemed unnaturally slow at tempo = 120. The accuracy of the Pattern 5 registrations was ± 1.74 msec. There were two usable versions of Pattern 5 registered, one with and one without guitar accompaniment. A third was rejected for analysis because it went far beyond the standard 15 measures.

As a starting point, the offsets were examined. Tables 43-45 give the summary statistics of the three distributions.

Table 43. Pattern 5, Version 1: Summary Statistics for the Overall, Bass and Snare Drum Distributions.

<u>Dist</u>	<u>Mode</u>	<u>Range</u>	<u>Standard Dev.</u>
Overall	1	16	3.817
Bass Dr.	-3	12	2.817
Snare Dr.	1	7	2.363

Table 44. Pattern 5, Version 1: Proportion of Time the Drum Hits Before, After or Simultaneous with High-hat.

<u>Drum</u>	<u>% Simultaneous</u>	<u>% Before</u>	<u>% After</u>
Bass	0	95.8	4.2
Snare	4.5	18.2	81.8

Again, one sees a tendency to hit the bass drum before the cymbal and the cymbal before the snare. Pattern 4 and 5 seem to share this tendency particularly, which may be due to the constant quarter-note figure performed on the cymbal in both. It is also possible that this manner of offsetting is a characteristic of Drummer B's style.

Pattern 5 offers an opportunity to compare accompanied versus unaccompanied drum performances. A striking point of comparison between the Pattern 5 solo and guitar-accompanied performances is in terms of tempo. The solo performance was played at tempo = 148, but the accompanied performance took

off at the much faster rate of tempo = 165! With such a wide variation in tempo, one wonders what other differences might be found in the accompanied performance.

Since there is only one accompanied performance that can be compared with an unaccompanied one in this study, no general statement can be made about any systematic differences in note durations between accompanied and unaccompanied drum performances. What can be said is that, in the case of Pattern 5, the correlation between the adjusted drum note durations of the two versions was .9952, indicating they are probably quite similar in terms of note durations.¹³

Before moving on to the broader comparison of linear regression equations, a look at the micro tempo maps of the two versions provides a visual comparison of note-level variation. Figure 23 (next page) shows the micro tempo maps for the accompanied and unaccompanied versions of Pattern 5 superimposed.

The most striking difference between the two maps is a wider range of micro tempo variation for the unaccompanied version. This difference warrants further study in following investigations, because synthesized drum rhythms that are intended for use with other instruments should be based on accompanied samples, if this difference turns out to be significant.

The regression equations also point up differences between the two versions. Table 45 gives the slopes and intercepts for them.

Table 45. Pattern 5: Slopes and Intercepts.

<u>Version</u>	<u>Slope</u>	<u>Intercept</u>
Unacc.	1.00284	-6.83
Acc.	.99640	82.48

The regression equation for the unaccompanied version is not remarkably different from others given in this investigation so far. It starts near the rational-mechanical tempo and slows down over its 15 measures. The accompanied version, however, has the slowest start and greatest long run increase of any performance analyzed yet.

13. For the sake of comparison, a casually selected sample of correlations from Drummer A's performances finds "r's" of .9982 between Pattern 1, Versions 4 and 5; .9970 for Pattern 3, Versions 1 and 3; and, .1024 for Pattern 2, Versions 1 and 2.

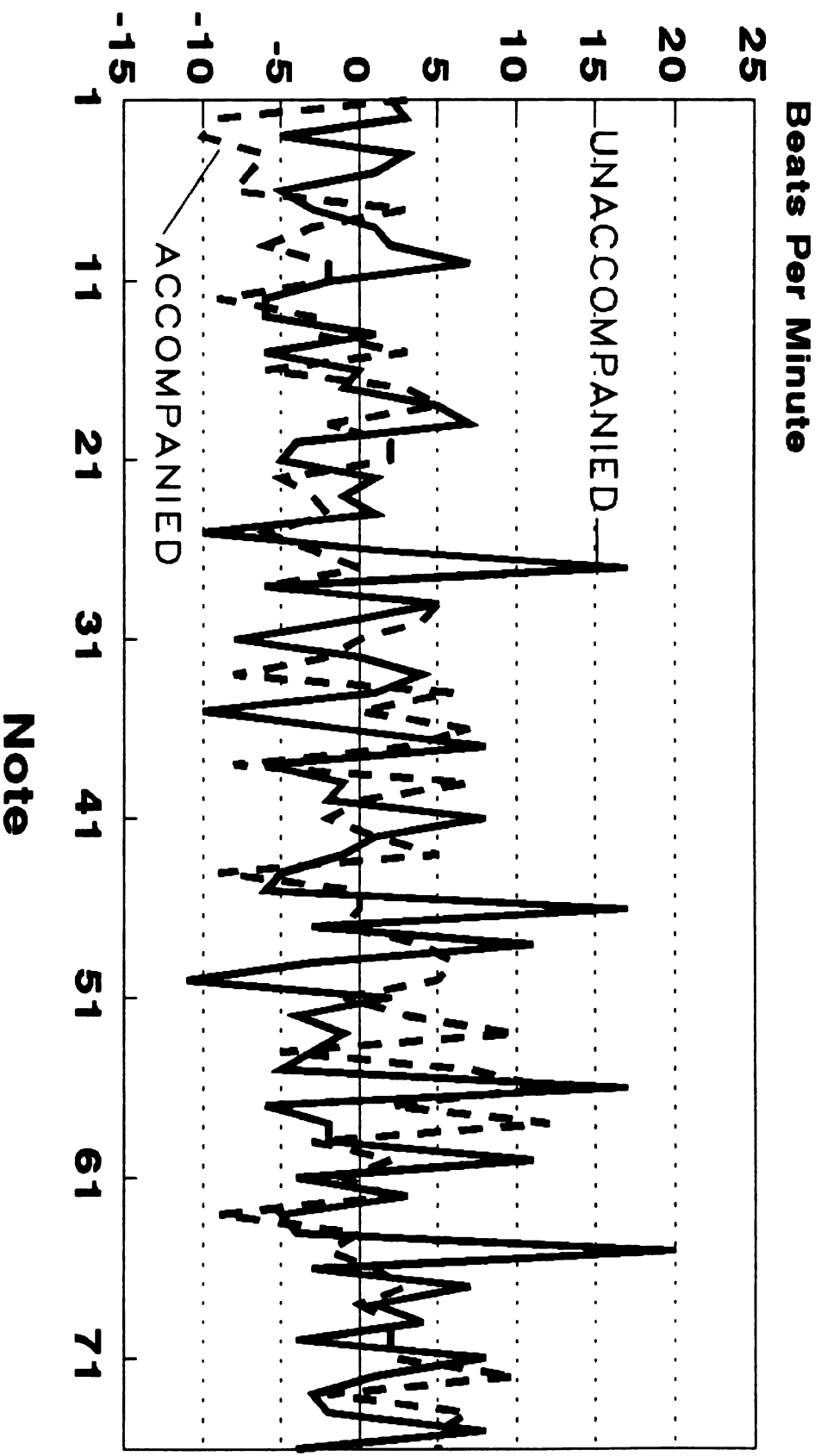


Figure 23. Pattern 5 Micro Tempo Maps

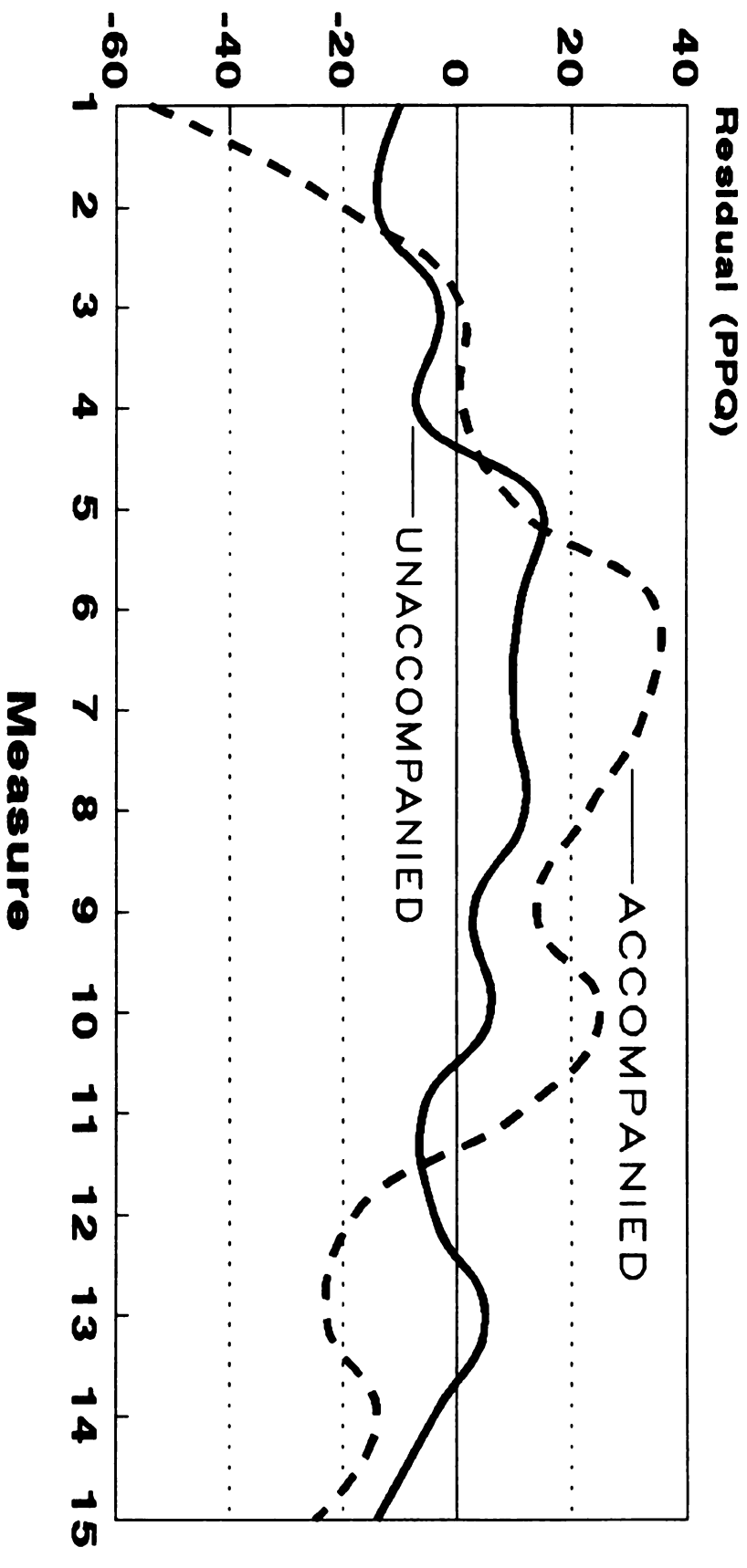


Figure 24. Residual Maps for Pattern 5

Its total duration is 6 PPQ long of the rational-mechanical norm, which means the slow start is very slightly over-compensated by the gradual tempo increase.

A last comparison is made with the residual maps of the two versions of Pattern 5. The maps shown in Figure 24 (previous page) are quite different. The unaccompanied version holds fairly close to the linear equation. The range of its residuals is 3 to 15. The accompanied version's map, however, has a wide arch. Its residual range, from 1 to 54, is also relatively wide. While a 54 PPQ error in the equation's prediction of the actual durations works out to less than a sixty-fourth note, still, the degree of non-linearity and the curvature of the residual map suggest that a non-linear equation might better describe the macro tempo of this version. Again, this points to a need for further study to compare accompanied and unaccompanied drum performances.

Pattern 6

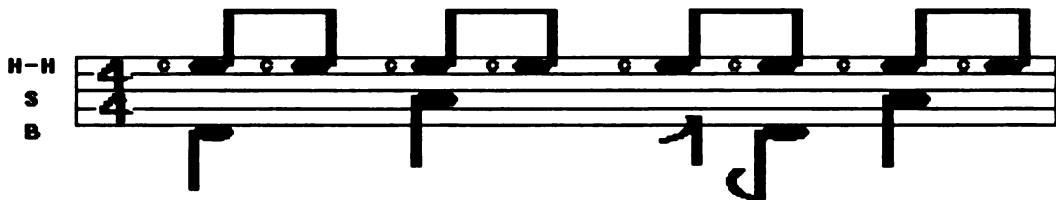


Figure 25. Notation for Pattern 6

Pattern 6 is a one measure pattern with an eighth-note rest on beat 3. It has a continuous eighth-note figure in the high-hat, as in Patterns 1, 2 and 3. Drummer B performed two versions of Pattern 6. The first look at Pattern 6 is in terms of offsets. Tables 46 and 47 give the summary statistics for the overall, bass and snare drum distributions.

Table 46. Pattern 6: Offset Summary Statistics for the Overall, Bass and Snare Drum Distributions.

<u>Dist.</u>	<u>Mode</u>	<u>Range</u>	<u>Standard Dev.</u>
Overall	-1	17	3.055
Bass Dr.	-4	17	3.543
Snare Dr.	1	9	1.996

Table 47. Pattern 6: Percentage of Simultaneous and Non-Simultaneous Hits.

Drum	<u>% Simultaneous</u>	<u>% Before</u>	<u>% After</u>
Bass	9.7	69.4	21
Snare	9.1	43.9	47

There is also found in Pattern 6 a tendency to hit the bass drum before the high-hat. However, the tendency observed above in Patterns 4 and 5 to have the high-hat hit before the snare is not very pronounced. A chi-square test of the Table 46 data proved significant ($p > .007$). This is due more to the bass drum offsets than to the snare. Also noteworthy is the low number of simultaneous hits in both the bass and snare drum distributions. One explanation for the offsetting differences between Pattern 6 and Patterns 4 and 5 is that the cymbal in the latter plays a quarter-note figure, while it plays an eighth-note figure in the former.

Table 48. Pattern 6: Slopes and Intercepts.

<u>Version</u>	<u>Slope</u>	<u>Intercept</u>
1	.99996	-2.51
2	1.00162	.88

Table 48 gives the regression slopes and intercepts for Pattern 6. The residual maps, shown in Figure 26 (next page), indicate that the Pattern 6 equations are quite accurate in predicting the actual note durations. The residual range for version 1 is from 1 to 8; for version 2, it is from 1 to 4 PPQ. Both versions start very near rational-mechanical tempo. Version 1 ends 4 PPQ short of rational-mechanical total duration, meaning that its tempo increase is extremely slight. Version 2 ends 28 PPQ long of rational-mechanical total duration. Thus, its macro tempo slows down in the long run from a near-normal start.

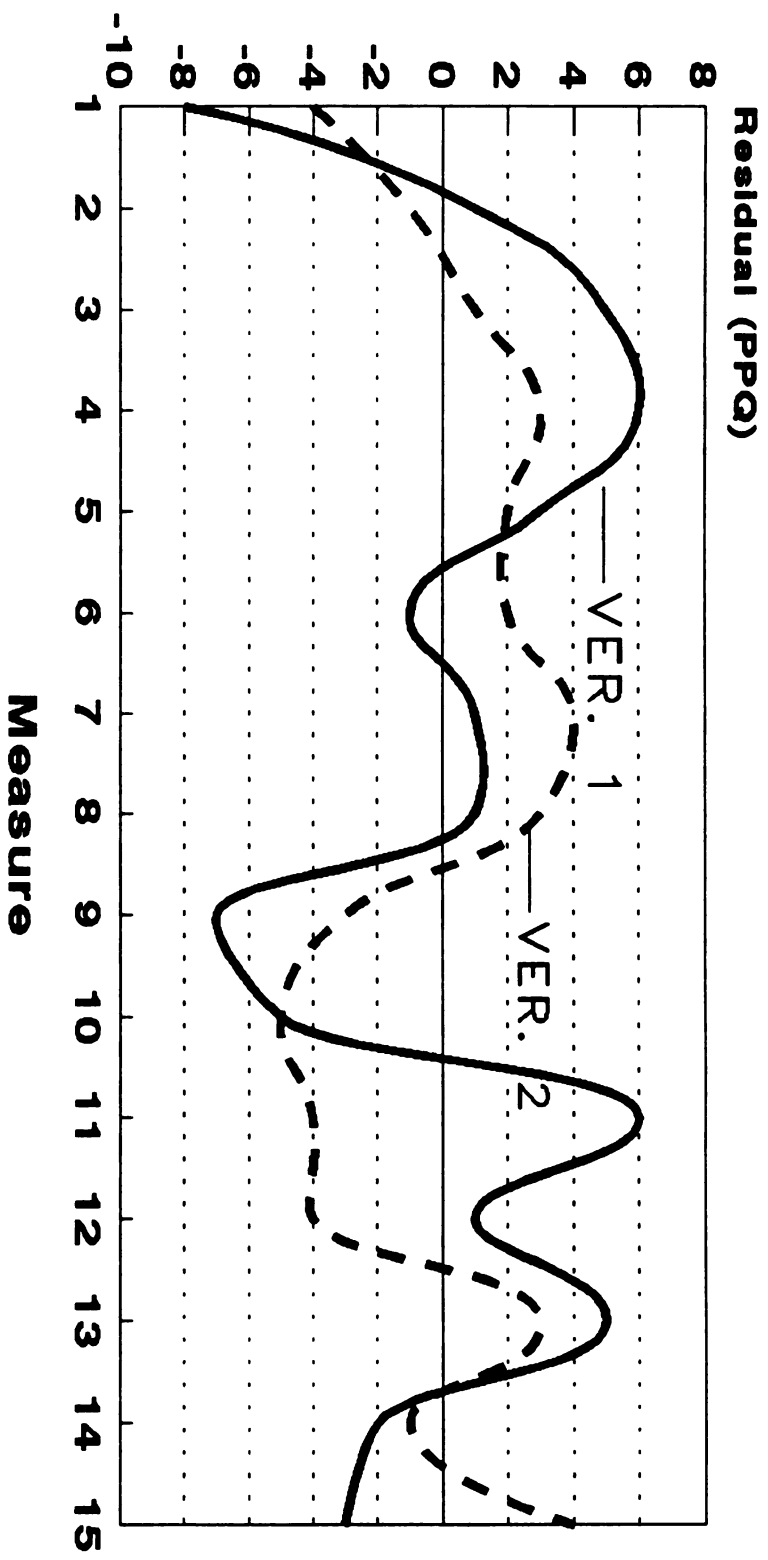


Figure 26. Residual Maps for Pattern 6

Discussion of Studio Experiments

The analysis presented above encompasses two different drummers, several different rock drum patterns, multiple versions of those patterns, and thousands of individual note durations. One of the difficulties in such an undertaking is figuring out how to make sense of such a large quantity of data. This study relied on the work of past researchers, the advice of musicians, and the author's own experience as a musician for guidance in the task. Though, as an exploratory study, the primary goal was to seek out the tangible ways in which a human drummer differs from a mechanical drummer, an attempt was made to analyze the data in a way that will help in producing better sounding synthesized drum rhythms.

In discussing the results of the studio experiment, something should first be said about the data collection method. This method was an experiment itself. Two different drum set-ups were used to obtain the data, one more natural for performance, the other more artificial. A comparison of these two set-ups is warranted to guide future work.

In the first studio session, acoustic drums were miked and fed through an analog signal-to-MIDI data converter box. Because most drummers play acoustic drums, this set-up provides the most representative data. The equipment used in the second session, a set of MIDI drum pads, is clearly less representative. One reason why is that the pads do not physically respond to the drummer's sticks as do acoustic drums. Watching Drummer B play, it was the author's judgment that this was a significant constraint on his ability to play naturally. Also, the dynamic response in a pad set up is purely a function of the MIDI velocity data, which has only a 128 step range. This means that the system is conforming the drummer's dynamics to its own limitations, further restraining natural play.

The use of an acoustic set, however, is not without its difficulties. The problem of cross-triggering is a formidable obstacle to smooth, reliable data collection. The advantage of using pads is that cross-triggering is quite easily controlled, and they are simpler to set up than an acoustic drum system. The good news is that the use of adhesive drum triggers, instead of microphones, as an input to the MIDI converter is likely to significantly reduce or eliminate cross triggering. Working with adhesive triggers requires some extra effort over a pad set-up, but it will ultimately provide better results and a more comfortable experience for most drummers.

In discussing the results of the data analysis, we begin at the note level. One thing that is strongly

evident, in all twenty registrations taken in the study, is that consecutive notes of equal rational-mechanical duration are least likely to have equal actual durations. In other words, when a drummer hits two eighth-notes in a row, the actual durations will follow a long-short or short-long pattern most often. This directly supports the findings of other researchers mentioned above.

While the actual durations of those two eighth notes are generally not going to be equal, the difference may not always be significant. In a given registration, as little as twelve percent of the consecutive differences were found to be greater than or equal to 10 msec. In other registrations, the figure was as high as forty-three percent. Furthermore, the percentages of what have been called "consequential differences" vary somewhat from pattern to pattern. Thus, they may be pattern-specific, to a certain extent.

These low percentages may actually be a boon in the synthesis of rock drum rhythms. The 10 msec threshold is here taken as an assumption, borrowed from previous research. If more systematically defined cut-off interval is known, the data from a registration can be reduced automatically according to it. More importantly, in synthesis of rock drum rhythms, whether by drum machine or sequencer, "humanizing" a pattern can be accomplished more efficiently, knowing only a certain percentage of consecutive notes of equal notational value need to be manipulated.

Similar to the note level phenomena, there is evidence of duration differences at the beat level. In Pattern 1, beat 4 in an average measure was made longer than beat 1 of the following measure. In Pattern 2, the average measure had beats 2 and 4 lengthened and beats 1 and 3 shortened. This suggests that *durational accents* may be occurring. This is also in line with the earlier research mentioned above, particularly that of Woodrow.

The interaction of the cymbal and the drums is a clear contrast from the rational-mechanical norm. Where in a notational sense a drum and cymbal were supposed to hit simultaneously, in actual performance, the drummer would most likely delay one somewhat behind the other. This held true for every pattern examined, except Pattern 2, where the snare quite often occurred simultaneously with the cymbal. Furthermore, the tendency to lead with either the cymbal or the drum was found to be drum-specific. If the the bass drum led the cymbal (as in Patterns 4-6), the cymbal would lead the snare. The chi-square analyses from Patterns 4 and 6 point out the trend convincingly. Finally, it seems that the range of the delay is generally smaller for the snare than for the bass drum.

Looking at the registrations in terms of macro tempo, there are several ways the drummers paced the 15 measure performances. Table 49 sums them up.

Table 49. Overall Macro tempo Behavior.

<u>Type of Start</u>	<u>Grad. Tempo Incr.</u>	<u>Grad. Tempo Decr.</u>	<u>Total</u>
Fast	10%	20%	30%
Normal*	30%	20%	50%
Slow	20%	0%	20%
	-----	-----	-----
TOTAL	60%	40%	100%

* A "Normal" start is one where the intercept, x , is $-10 > x > 10$ PPQ. "Fast" start is where $x \leq -10$; a "Slow" start is when $x \geq 10$.

Of the six possibilities for macro tempo variation, no one has a clear majority. However, there is a 60/40 majority in favor of gradual tempo increases. Normal starts account for fifty percent of registrations; of the other fifty, fast starts have the majority.

There is not a clear trend in the macro tempo variations. Therefore, it is not possible from this data to come up with a general statement on macro tempo change one can apply in synthesizing rhythms. Suffice it to say in every registration there was a macro tempo change of some sort. Though no statement can be made about how macro tempo was generally performed in the registrations, it is enough to have found evidence of macro tempo change, and a compact means of describing it (whether by linear or quadratic model). To show clearly that it is prevalent points to yet another difference between a mechanical and a human performance.

Another point of contrast, though not extensively explored, is that accompanied drum performances appear to differ from unaccompanied ones. Evidence of this is found in the regression equations, micro tempo and residual maps for Pattern 5. The guitar-accompanied version had the slowest start, fastest overall tempo increase and widest range of residual error for any registration in the study. Its residual map also had the greatest arch. However, the magnitude of variation of its micro tempo map was noticeably less than for the unaccompanied version. This suggests that playing with a guitarist had opposite effects on micro and macro tempo, an interesting finding worthy of further study.

A final comment on the results of the studio experiments. It is important to keep perspective on the differences observed between the actual note durations performed by the drummers and the rational-mechanical norm. There are doubtless many reasons other than note duration variance why listeners can tell the difference between human and mechanical rhythm performances. Some of these other reasons may be: dynamics, timbre, what combination of other instruments (if any) are combined with the drum sounds, musical preference of the listener, and listener attitude toward synthesized music. A combination of the performance factors and the psychological predisposition of the audience will determine in each listener whether he or she prefers a human over a mechanical performance or is indifferent. While analysis of the 20 registrations taken in this study has uncovered definite patterns of variance from the rational-mechanical norm, it remains to be seen how important this variance is, as captured by the sequencer, to the audience. One must test with an audience the assumption that durational differences of the order found in the above analysis contribute significantly to the listener's enjoyment of a musical rhythm.

AUDIENCE TEST

Method

An interesting question that arises out of the above analysis is: how sensitive is an audience to systematic variance in these types of rhythms? The set-up makes it convenient to produce versions of a given rhythm with systematic, random or no variance from the rational-mechanical norm. Thus, it is possible to manipulate the type of variance applied to a pattern and test the result on an audience.

Master Tracks Pro (TM) has a humanize function, as stated above. It also has a "quantize" function that will eliminate any variance from the rational-mechanical norm. The solo drum patterns were quantized and humanized to produce different versions of the various patterns for use in the audience test.

The audience test was divided into three parts. Part I dealt with audience preference for human, computer generated and humanized performances of drum beats. There were eight pairs of drum patterns, each containing alternate versions of the same pattern. Five of the pairs were made up of a human performance (collected from one of the studio sessions) and a quantized version of the pattern. These included two accompanied performances created by adding MIDI bass notes and harmony to the drum sequences. The bass and harmony parts were recorded into the sequencer in real time with a MIDI guitar. They were spare, notationally identical parts, intended not to bury the drums sounds, but to place them in the context of other instruments. Thus, it was possible to see whether adding instrumentation to a drum pattern affects the listener's ability to perceive differences in systematic variance in the drum pattern.

In addition to the five pairs used for comparison of human versus quantized performances, three additional pairs were included in Part I. One pair consisted of a humanized and a quantized performance. Another contained a tape recorded human drum performance (using synthesized drum tones) and the computer-registered version of the same performance. Finally, one pair was made up of duplicate versions of the same quantized rhythm, as a control.

In all eight pattern sets making up Part I, the same synthesized drum sounds were used, to control timbre as an intervening variable. Also, the velocity values between pairs were set equal. This focused the tests in Part I exclusively on durational differences.

The listeners were asked to respond for each pair, "which version sounds best to you, or whether the versions sound the same." Further instructions were given for the respondents to use their "feelings and sense of rhythm," and that it was "not necessary to think it over," because "there is no right or wrong answer."

Part II was designed to test how well listeners can perceive small differences in the onsets of two notes. In three separate tests, a high-hat and bass drum tone were spaced fifteen, ten and four milliseconds apart, respectively. The subjects were asked to identify, in each case, whether the bass drum or the high-hat struck first, or whether they struck simultaneously.

Part III tested listener ability to perceive durational differences between consecutive pairs of notes. Groups of bass drum notes were generated containing pairs of notes separated by varying time intervals. The pairs and their spacings were as follows: 1) 104 vs. 52 msec; 2) 52 vs. 21 msec; 3) 21 vs. 10 msec; 4) 135 vs. 125 msec. This test was an attempt to verify the 10 msec threshold for human perception of durational differences given by Seashore.

The sound stimuli were assembled on a digital audio tape, which was controlled by the author in each of the three test sessions. The tape was stopped between each new version or test, and the upcoming event announced, prior to playing the event for the audience. An oral synopsis of the questionnaire's written instructions (see questionnaire in appendix) was given before each Part. Questions were handled at that time and the test proceeded. Subjects were allowed to hear each event only once. The participants were undergraduate students in a basic audio production class at Michigan State University.

Results

The audience test was conducted in three sessions, over a two day period. The sample was made up of 28 undergraduate students in a basic audio production class. Demographically, the sample consisted of fourteen female, and fourteen male subjects, ranging in age from 19 to 25. The average age was 21 years.

Seventy-one percent of the subjects had some sort of past musical training. The range of training in years was one to fourteen, with an average of four. Twenty-nine percent of the subjects described themselves as currently playing one or more instruments. These had an average of two years' experience on their current instruments.

The subjects claimed they listened to an average of four hours of music per day. Current players and non-players listened to about the same amount of music per day. Collectively, the subjects thought it fairly important to listen to their music on a good quality stereo system (average was three on a one to seven scale, one meaning "very important").

As a whole, the subjects expressed slight disapproval of the use of drum machines in music they liked. The average score was five on a one to seven scale, seven meaning "strongly disapprove." However, there was a significant difference (t-test, $p > .01$) in approval of drum machine use between current players and non-players. The mean for current players was six, and that for non-players was four. This means the current players profess stronger disapproval than non-players. It seems disapproval of drum machine use may increase somewhat with years of experience for current players, as well. Years of experience has a .36 correlation with disapproval of drum machine use.

Current players were as likely as non-players to incorrectly label different versions of the same pattern as having no difference. Nor did years of experience seem to improve the current musicians' acuity. The overall average for erroneously identifying a pair of alternate versions (in Part I) as having no difference was thirty-eight percent.

The Part I comparisons of alternate fifteen measure drum performances will be reported separately, by Pattern. They will be presented in the order they appeared on the questionnaire.

1. Pattern 3, Version 2/Quantize

Twenty-one percent of the subjects perceived no difference between the human and quantized performances. Fifty-four percent thought the human version sounded best, and twenty-five percent thought the quantized version sounded best. Of those who perceived a difference between the versions, sixty-eight percent preferred the human performance. The human version was given first.

2. Pattern 5, Version 1/Quantize

Twenty-eight percent reported no difference. Fifty percent preferred the quantized version; twenty-one percent preferred the human version. Seventy percent of those perceiving a difference preferred the quantize over the human version. The quantize version was given first.

3. Control (Quantize/Quantize)

Fifty-seven percent perceived no difference. Twenty-nine percent claimed the first version sounded best; fourteen percent chose the second version as best.

4. Pattern 2, Version 4/Quantize

Thirty-five percent of the subjects reported no difference between the two versions. Forty-six percent chose the human performance as best sounding, while seven percent chose the quantize version. Of those perceiving a difference, seventy-two percent chose the human version. The human version was given first.

5. Quantize/Humanize

Sixty-four percent saw no difference. Twenty-nine percent chose the humanized version, leaving seven percent for the quantized version. Eighty percent of those perceiving a difference chose the humanized version. The humanized version was given first.

6. Computer/Tape Recorded

Only four percent of the respondents perceived no difference between the sequencer-generated and tape recorded versions of this performance. Sixty-one percent preferred the computer-version over thirty-six percent for the tape recorded version. The computer version was given first.

7. Pattern 4, Version 2/Quantize (accompanied)

Twenty-nine percent perceived no difference between the drum performances in the two accompanied performances. Thirty-nine percent preferred the quantized version. Thirty-two percent chose the human version as sounding best. Fifty-five percent of those perceiving a difference preferred the quantized version, which was given second.

8. Pattern 6, Version 2/Quantize (accompanied)

Forty-three percent saw no difference. Thirty-six percent reported the quantize version sounded best to them, while twenty-one percent chose the human version as best. Of those perceiving a difference, sixty-three percent preferred the quantized version. The quantized version was given first.

The results from Part II, the portion of the test dealing with offsets, will be given next. As with Part I, they will be presented in the order in which they appeared on the questionnaire.

1. 15 Msec Offset

The bass drum hit first. Forty-six percent were correct in identifying this; fourteen percent thought the high-hat hit first, and thirty-nine percent thought they hit simultaneously.

2. 10 Msec Offset

Again, the bass drum hit first. Thirty-two percent were correct; thirty-six percent thought the cymbal hit first, and thirty-two percent thought the hits were simultaneous.

3. 4 Msec Offset

The high-hat hit first. Thirty-two percent were correct; twenty-nine percent said the drum hit first, and thirty-nine percent thought the hits were simultaneous.

Overall, eighty-nine percent of the respondents got one or two of the offsets correct. And current players were not relatively more accurate than non-players.

Part III, the duration differences tests, are given last. They are presented in the order they appeared on the questionnaire.

1. 50 Msec Difference

The second pair was spaced farther apart in time. Eighty-nine percent of the respondents were correct in naming the second pair as spaced farther apart.

2. 30 Msec Difference

The second pair was spaced farther apart. Eighty-nine percent of the subjects correctly identified this.

3. 10 Msec Difference

The first pair was spaced farther apart. Seventy-nine percent were correct in identifying this.

4. 10 Msec Difference (135 vs. 125 msec)

This test was designed to see if a small (10 msec) duration difference can be detected when it is between notes of longer duration, such as two sixteenth notes. The first pair was spaced farther apart. Seventy-five percent of the respondents correctly identified this.

Overall, sixty-four percent of the subjects correctly identified all of the duration differences.

Discussion

The audience test results offer no clear conclusion as to whether listeners generally prefer human over quantized drum rhythms, or, vice versa. In three of the five tests where listeners were asked to choose which "sounded best" between a human or quantized version of the same rhythm, they chose the quantized version. However, this set of five includes the accompanied versions. For the unaccompanied pairs, listeners chose the human versions two out three times.

There is also a suspicious pattern in the responses for Part I. In seven of the eight tests, the listeners chose the first version, whether it was a human or quantized performance. In the control test, where two identical quantized versions were paired together, while the majority (57%) chose "no difference," those perceiving a difference favored the first version 2:1 over the second. This suggests there may be a problem with the format of Part I. The pattern of "first version" responses could have occurred by chance or may represent true listener preference, but it is also possible that, when listeners perceive a difference between the two versions, there is something about the format predisposes them to favoring the first version.

Whether the test format biases listeners toward choosing the first version or not, the results of Part I do not confirm listener preference for either human or quantized performances. An index was created to represent listener preference for human over quantized performances. The index ranges from minus five to plus five, plus five meaning a listener always chose the human performance, minus five meaning the person always chose the quantized version. A score of plus five or minus five necessarily means the individual could always perceive a difference between the two versions. The mean value of this index was zero. Thus, the results from this sample indicate a draw in preference for human vs. quantized performances.

There are several possible explanations for this equivocal result. One possibility is that duration differences alone do not bias listeners strongly enough for or against quantized drum performance. Timbre and dynamics are the two other variables that listeners can use to distinguish human from machine-generated drum performances. In this case, the listeners may have lacked sufficient information to make a clear choice, and may have felt unsure or indifferent about their decisions. Another possibility

is that listeners tuned in on the two drummers' different playing styles, or on duration differences brought about by the use of an acoustic kit versus a set of MIDI pads in registering the performances. The listeners chose the human performance for Drummer A two out of two times, and the quantized performance for Drummer B three out of three times.¹⁴

Though the results point out no clear preference for human or quantized performances of the drum patterns, other conclusions can be drawn from the data. First, in all but two cases (Control and Quantize/Humanize), the majority of listeners stated they perceived a difference between the two versions. The "no difference" proportion of the audience for these tests ranged from four to forty-three percent. The greatest number of "no difference" responses occurred for those tests not including a human performance. The listeners were better able to distinguish between the two versions when the pair included a human performance.

Second, an ironic finding has to do with the subjects describing themselves as currently playing an instrument. Recall that this group was significantly more opposed to the use of drum machines in music they liked. Yet, these subjects displayed a tendency to choose quantized over human performances. The correlation between current player status and preference for quantized versions is .36. There was a weaker correlation (.25) between these current players' years of experience and preference for quantized versions. It would be interesting to explore this tendency further. Perhaps, the current musicians believe that drum machines are *less precise* than human drummers, and interpreted any variance they observed between alternate versions as machine error.

Two tests were included in Part I which did not seek to compare human with quantized performances. These were the humanize/quantize comparison and the computer registered/tape recorded comparison. In the humanize/quantize test, the majority, or, sixty-four percent of the listeners found no difference between the two versions. Of those who found a difference, eighty percent thought the humanized version sounded best. When it "humanized" the performance, the computer was given a 21 msec range in which to randomize note start times. This range was selected based on the average range found in the various human performances. Thus, most of those who could perceive a difference preferred the humanized version with a randomization range of 21 msec. However, the majority of listeners perceived no difference at all.

14. This comparison is somewhat tenuous, because two of Drummer B's three versions are accompanied, while none of Drummer A's are.

The computer registered/tape recorded test had a curious result. For it, the proportion of listeners reporting "no difference" (4%) was lower than for any other test in Part I. Sixty-three percent of those perceiving a difference chose the computer registered version as best sounding. The rationale behind this test was to see if listeners could discern the difference between the performance as justified to the 240 PPQ resolution of the sequencer, and the same performance passing through the sequencer unjustified and being recorded simultaneously to analog tape. The listeners may have been judging sound quality, in this case, instead of rhythm quality. In terms of sound quality, the computer generated version would be superior, since it was transferred to the DAT as a first generation analog recording. The tape recorded version, however, was transferred to the DAT as a third generation analog recording, which implies sound quality inferior to the computer generated version.¹⁵

Another explanation for the results of the computer generated/tape recorded test is that a rhythmic difference was detectable, but the fact that the computer generated version was given first somehow predisposed listeners to chose it. Whatever version the majority of listeners chose for this test, one thing stands out brightly. That is, the four percent figure for "no difference" responses. This test used the drum track from a guitar accompanied performance (Drummer B, Session 2) as the pattern. What distinguishes the pattern is that the drummer was allowed to make it up himself and improvise at will. As a result, there were many more fills included. The greater volume of notes per unit of time in this performance may have provided the audience with better cues (certainly, more information) with which to chose the best sounding version.

The three tests in Part II were designed to check whether listeners could detect offsets between drum and high-hat start times. In every case, the majority thought there was an offset, though correctness of determining which sound hit first varied. The highest proportion of correct answers were for the 15 msec offset, not surprisingly. As the offset diminished from 15 to 10 msec fewer correct answers were given. Interestingly, slightly more persons thought the high-hat hit before the drum (the reverse was true) than answered "simultaneous hit". Perhaps, this is because the sound of an offset is easier to detect than the actual order of the hits. As the offset diminished from 10 to 4 msec, more persons thought the hits were simultaneous

15. However, each time the tape recorded version was transferred, from the 24 track master, to a half track machine, to another half track machine, the recording speed was 15 ips. Technically, the tape recorded version was inferior sounding, however, subjectively, this may not have been the case.

than thought there was an offset. This shows that some listeners can perceive offsets when they are focused on them, and particularly if they are in the 15 msec range. However, the importance of offsets to listener preference of a drum performance is uncertain. And most offsets found in the registrations of this study fall in the 0-10 msec range.

It might be better, when focusing on the importance of offsets to the rhythmic quality of a drum performance, to ask listeners, simply, whether a drum and cymbal hit with an offset sounds different than an simultaneous hit, and not ask what the order is. Clearly, the order is not what is important, rather, whether or not there is a perceptible aural effect.

Part III dealt with duration differences and produced the clearest results in the audience test. Eighty-nine percent of the listeners were able to detect 50 and 30 msec duration differences between bass drum hits. Seventy-nine percent were able to detect a 10 msec difference. Finally, seventy-five percent were able to detect a 10 msec difference between bass drum hits spaced over 100 msec apart.

The Part III results show most listeners as having a fairly keen sense for duration differences, down to 10 msec. How well they can detect small duration differences between notes beyond approximately 150 msec is an important question to answer in further studies. A quarter note at tempo = 120 has a duration of about 500 msec. A long-short orientation of quarter notes having respective durations of 510 and 500 msec may well be inconsequential to a listener. However, a long-short orientation of sixteenth notes at that same tempo, with respective durations of 125 and 115 msec, is likely to be noticeable. If this reasoning holds true, it may explain why an average of thirty-one percent of the subjects found no difference between the human and quantized versions across the five relevant tests. The patterns were generally made up of quarter or eighth notes, with only two to four sixteenth notes included at measure eight.

Support for this theory is found in the scientific literature on the *Time Sense*, which dates back as far as 1864. Through various means the researchers have sought to establish the precise limits of human perception of duration differences. Together these experiments establish that human beings have a variable threshold for perception of duration differences. This means that, where duration differences are proportionally equal, the absolute durations of the tones being compared has an effect on how accurately listeners can judge the difference in duration. Between roughly 100 and 200 msec, perception is keenest. Between 200 and 1000 msec, perception drops off significantly; at

levels below 50 msec or above 1 second, it is poorest (Michon, 1964).¹⁶

The implication of this variable sensitivity to the present audience test is that the majority of note durations in the test stimuli fall 250 to 500 msec range, a region of diminished sensitivity to duration differences, according to the research literature. This may explain why an average of thirty-one percent of listeners reported no difference between human and quantized versions.

In cases where quantized versions were preferred over human, the presence of "perceptual accents" may have been a factor. According to the following excerpt from Haydon, a perceptual accent "is an accentual effect produced when, in the course of a phrase, a rest occurs on a strong beat. A similar effect is often produced by a syncopation, that is, when no note is 'struck' on the strong beat. The so-called subjective accent is one 'read into' a perfectly even series of pulses." (Haydon 1941, pp. 164-165)

Such perceptual accents can be felt in the absence of any systematic variance from the rational mechanical norm. Patterns 3, 4, 5 and 6 contain syncopated beats where listeners might experience perceptual accents. Of these four patterns, quantized versions were preferred 3:1, though by narrow margins. In Pattern 2, which did not contain any syncopated beats, listeners preferred the human version.

Summing up the audience test, there seems no justification to reject drum machines generating repetitive beats according to the rational-mechanical norm as patently unsatisfying to an audience. The audience members of this study were not convincingly capable of distinguishing between human and quantized performances. And when they perceived a difference, they were equally as likely to choose a quantized as they were a human rendition of a repetitive pattern as best sounding. This equivocal result may be due to the fact that most note durations in the test stimuli fell in a region where listeners have diminished capacity to detect duration differences, and because most of the patterns contain syncopated beats where listeners might have perceived perceptual accents in the quantized versions.

This aside, it is important to note that typical human rock drum performances are peppered with fills made up of short notes, and, in general, have notes varying in intensity and timbre. The duration tests in Part III and the scientific literature show that faster tempos and a

16. This does not necessarily contradict Seashore's 10 msec threshold. The 10 msec figure for "a very fine musical ear" (Seashore, 1938) may still hold, while being dependent on the absolute durations of the tones being compared.

greater volume of short duration notes would probably put the differences between human and quantized performances in starker contrast. Future tests should allow the drummers more freedom to play as they would normally, not force them to mimic a drum machine; they should also be allowed to improvise at will. A better test would match quantized versions of those freer performances with their human counterparts. This way, the test would challenge the drum machine to match (as programmed according to the rational-mechanical norm) the human performance and the audience would judge its success.

Producers weighing the merits of using a human drummer versus a machine are advised that, for a repetitive pattern, even discriminating listeners, such as, musicians, will not definitely perceive the difference by duration differences alone. Doubtless, timbre and relative note intensity are important cues that tip off an audience to use of a drum machine. Even then, not every audience member will be strongly prejudiced against its use. Not surprisingly, musicians will probably be most offended by it, as they reported to be in this sample.

If the intent is to disguise the use of the drum machine or sequenced drum performance, however, and that performance contains a good deal of fills and other short duration notes, the producer is advised to edit the sequence with careful attention to those areas. This should be done with the knowledge that listeners can perceive relatively small duration differences between short notes.

RECOMMENDATIONS FOR MUSICIANS AND FUTURE RESEARCHERS

For Musicians

One of the goals of this study was to arrive at some applicable knowledge that musicians might use to make their drum sequences sound more human. While the data from the studio experiments provide only information on note duration variance, yet, a few recommendations may be made based on the available data.

From the outset, it should be kept in mind that the audience test did not confirm that listeners routinely discern even actual human performances of repetitive patterns from rational-mechanical likenesses of them. And even when they can, it is not at all certain they will prefer the human performance over the mechanical. Putting aside reservations about how representative the sample is or possible problems with the method, these are the basic conclusions of the audience test.

This said, if a musician decides his drum sequence sounds too mechanical, here are some tips on how it might be made to sound better.

Tip 1. Play and Analyze

If you are using a sequencer or a drum machine that will record (and not quantize) real time rhythmic input, and the sequencer or drum machine will allow you to ascertain the durations of individual notes, first, play the basic pattern into the device several times. Then, take a look at the note durations and see if any patterns emerge. It is likely some will. This procedure was followed in the studio experiment pretest to good effect. In that trial, it became clear that a grouping of consecutive eighth notes was consistently being performed long-short, in terms of actual duration. If something similar emerges in your data, it is probably no accident. What you are seeing likely is *systematic variance*.

Where note durations are short, this type of variance probably has a more noticeable effect. Therefore, in copying out the pattern to phrase or song length, try incorporating the systematic variance as a means of improving the rhythmic feel of the pattern.

Look also for systematic variance in terms of note intensity or MIDI velocity. Accented notes may have

consistently greater velocity values than their surrounding notes. Check for systematic intensity variance in fills and on cymbal figures, as well. These patterns of velocity or intensity variation can be incorporated when copying out the pattern to phrase or song length.

Tip 2. Try a Slight Tempo Change

A convenient way to add some tension to a bland sounding repetitive pattern is to apply a slight tempo change. An overall change in the range of three to four beats per minute, perhaps with a beginning tempo slightly above or below average tempo, will probably not be noticed as a speed up or slow down as such, but may alleviate some of the blandness and provide a sense the rhythm is "going somewhere." It may be necessary to do some editing at the note level, if, say, a speedup becomes too pronounced near the end of a phrase or song, or if a slow start sounds a bit sluggish. If this happens reduce the linearity of the tempo change with individual note editing. Reducing the duration of consecutive notes (i.e. bringing their start times closer together) will have the effect of speeding up the tempo; increasing consecutive note durations will slow tempo down. Use your ear to find the spot and tell you what needs to be done.

Tip 3. Humanize After Editing Structurally

Assuming that you have found some systematic variance and incorporated it with or without a slight tempo change, it might be time to add some random variance as a final touch. Set the amount of randomization with the knowledge that you have already put some variance in, and what you are doing is attempting to *reduce the redundancy of that systematic variance*. A randomization amount of less than 10 msec might not be noticeable. Then again, if it adds to something in the wrong way, 10 msec might make the sequence sound odd in spots. Experiment with different amounts of random variance, and be prepared to do a little individual note editing if certain parts sound out of whack.

Tip 4. Try Offsets Between Drum and Cymbal Sounds

According to the data in this study, simultaneous drum and cymbal hits are comparatively rare in repetitive drum patterns. Another way of "humanizing" a sequence therefore is to offset the start times of the cymbals to the drums. The tendency of the two drummers in this study was to hit the bass drum before the cymbal and the cymbal before the snare, or vice-versa. Also, the offset interval was generally larger for the bass drum offsets than for the

snare. A range of 5-10 msec for the bass drum and 0-6 msec for the snare was generally the case in the above data.

Tip 5. Listen To Human Drummers

If what you are trying to do is benefit from the economy of a sequencer or drum box, without drawing attention to the fact that your drum line is mechanically generated, you should try to write a drum part that makes sense from a compositional standpoint. That is, one which a human drummer might typically play. There are certain rhythmic patterns, which a drum machine can do all day, that would exhaust a human drummer in minutes. There are other things a drum machine can do that a human drummer would have to have three arms and three feet to match. The idea is that if your sequencer or drum machine is performing super-human feats, it is not sounding human. This defeats your purpose. The way to guard against this is to listen to human drummers and take mental notes of what they do at key points, such as, phrase points, fills, song sectional transitions, etc. Then, write accordingly.

Tip 6. Slower Tempo Settings Affect Real-Time Resolution

This last tip involves technical considerations and beat resolution. It is important to realize, particularly when working with low resolution (120 PPQ or less) sequencers and drum machines, that the *real time resolution* of the unit decreases as the tempo setting is lowered. This is in spite of the fact that the beat resolution remains constant. This is because sequencers and drum machines typically alter tempo by speeding up or slowing down their event clocks, thus, increasing or reducing the real time value of one beat resolution unit, or, PPQ. While this is a convenient way to engineer a sequencing device, it causes rhythmic performance to vary with the tempo setting.

Take, for example, a sequencer with 120 PPQ resolution. At tempo = 100, the sequencer is capable of manipulating note durations in increments of 5 msec. Since 5 msec falls below the 10 msec threshold for human perception of duration differences, this level of accuracy may prove acceptable. However, when the tempo setting is lowered to, say, 60, the sequencer can only manipulate note durations in increments of 8 msec.

This reduction may or may not present a problem, depending on the note values included in the sequence and how the sequencer is being used. For instance, the real time value of a quarter note in the above example, tempo = 60, is 960 msec, or, almost one second. At that length, it

is doubtful listeners can discern between, say, a 120 PPQ note duration (960 msec) and a 121 PPQ duration (968 msec). However, at shorter note durations, for instance, down around a thirty-second note, where the real-time rational-mechanical duration equals 120 msec, the sequencer may seem less responsive or stiff to the musician using it. In other words, what is played in and what comes back out sound like two different things, rhythmically. Also, the shorter note durations put this lack of responsiveness closer to the listeners' threshold of perception, as well.

Finally, a word of caution when recording difficult parts into a sequencer at slow tempo, thinking that they will "sound the same, only faster" when the tempo setting is raised. This feature is touted as a great benefit of using sequencers over tape recorders, because speeding up a tape recording changes the pitch of the recording, while this does not happen with a sequencer. Unfortunately, recording a part at tempo setting 60 on a 120 PPQ beat resolution sequencer and speeding it up to tempo setting 100 has the same effect as recording the part at tempo setting 100 on a 60 PPQ resolution sequencer. Figure 27 illustrates this point:

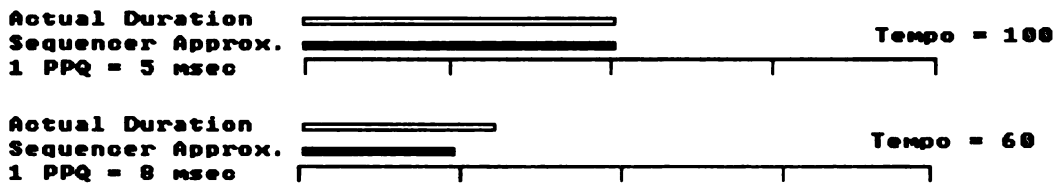


Figure 27. 10 Msec Note Recorded at Tempos of 100 and 60

The actual (real-time) duration of the hypothetical note is 10 msec. At 120 PPQ beat resolution, if it is recorded into the sequencer at tempo setting 100, it is given a 2 PPQ duration by the sequencer. On the other hand, if the note is recorded at tempo setting 60, the note is given a 1 PPQ duration by the sequencer, because its actual length is within 8 to 16 msec, and is closer to 8 than 16. When this note, recorded at tempo setting 60, is played back at tempo setting 100, its 1 PPQ designation will make the sequencer give it only a 5 msec actual duration. The difference between the 10 msec actual duration it should get, and the 5 msec duration it does get, is the error caused by recording it at a slow tempo and speeding it up. Though the actual duration at tempo setting 100 is supposed to be less than 10 msec, proportionally, the note's actual duration is incorrect to its original performance.

For Future Research

For researchers looking into the rhythmic character of rock drum performances, there are some recommendations coming out of this study, too. The first has to do with using a MIDI sequencer to register drum rhythms.

It has been mentioned above that it is the author's judgment an acoustic drum kit and adhesible triggers will produce the most representative data. This covers the controller end of the system. The other end is the receptacle of the MIDI data, that is, the sequencer. There are several advantages of using a sequencer over some analog method to capture drum rhythms for analysis. To begin with, note start times are automatically determined and registered in the act of recording the performance. Other researchers (i.e. Bengtsson and Gabrielsson) have gleaned their data from strip chart recordings of analog signals. A drawback to this method is that one must manually ascertain the start times and extract the note durations from the wave forms on the strip chart. Not only is this a laborious, tedious undertaking, but the accuracy of the registrations does not lend itself to tight estimation. This is because different instruments have different attack envelopes. Finding the exact onset of a cymbal sound may be more or less difficult than finding that of a bass drum. The attack envelope (consider a ride cymbal) can even vary depending on where and how the instrument is struck.

A MIDI setup removes the step of manually finding the note onsets from a strip chart. Also, the window of error is about equal to the real-time value of 1 PPQ. Moreover, this window is grounded on the stability of a computer's crystal-based clock. The actual error of a sequencer registration may be estimated based on the clock's variance, the sequencer's beat resolution, and its tempo setting. For most registrations in this study, the window was around 2 msec. Earlier studies, using other methods, estimate registration errors in the range of 10 msec.

Registering with a sequencer has advantages in terms of accuracy and automation of start time data, however, the author has yet to come across one designed exclusively for musical rhythm research. In practice, the sequencer, designed to be efficient in creating music, is not always so when it is being used to study music. The graphical displays showing note placement and duration (i.e., the typical "piano roll" display) work far better with quantized note data than with human performances. Gradual tempo changes in human performances often make the static measure boundaries in these displays useless and confusing when interpreting the registration. This and other problems are a price the researcher must pay for the editing, data

manipulation and storage capabilities that are the benefits of working with a MIDI sequencer.

If the researcher elects to register musical rhythms with a MIDI sequencer, it is best to choose one that supports the Standard MIDI File. It is relatively quick and easy to write a short program that will convert MIDI note start time data from a SMF to note durations, measured either in PPQ or milliseconds. With this capability, the registered performance can be taken out of the sequencer and put into a spreadsheet or statistical package with a minimum of time and effort, using an ASCII file. The alternative to this is computing durations manually from individual note start times, a time-consuming, tedious process similar to that required with the analog method mentioned above.

Another piece of advice is to discard the notion that the sequencer's graphic note display (if it has one) will be of much use. Whatever tempo the researcher wishes the musician to play at, the sequencer should be run at maximum tempo setting. This will minimize the time value of 1 PPQ, and maximize the accuracy of the registration. Using MasterTracks Pro, the accuracy of the registrations in this study could have all been around 1 msec, had it been known the graphic display would be of so little use.

Finally, some observations on the structure of future studies in this area. Future studies may wish to reverse the order used in this one, placing the audience testing portion first and the data analysis last. While this seems odd, there is a good reason for it. It is better to first find out what an audience prefers, and then analyze that, than it is to analyze something not knowing whether the audience prefers or is indifferent to it. In the former case, the fruits of analysis can be used to synthesize rhythms that have probable audience approval; in the latter case, what seem substantial findings in the analysis stage may prove to be of questionable value after the audience testing.

As for the audience testing, resources should be set aside to spend considerable time pretesting and, perhaps, developing innovative ways of delivering test rhythms to the audience. There may be a problem in asking the listener to discern between two versions of a rhythm presented in series. Some sort of parallel presentation, allowing the individual to select in process which version he wishes to hear, may prove more effectual. Also, more work needs to be done to determine which variance in a rhythmic performance is of significant consequence to the listener, and which is not. Better audience testing will eventually produce better synthesized rhythms.

Appendix A

Sample Audience Questionnaire

MUSICAL RHYTHM STUDY

Orientation

Your participation in this study is voluntary and optional. By filling out the questionnaire, you indicate your agreement to participate in the study. Your answers will remain anonymous, not linked to your name in any way.

You are about to participate in a study that concerns musical rhythm. There are three parts to this experiment. In part I, you will hear pairs of short drum rhythms. Each pair contains alternate versions of the same rhythm. You will be asked to answer which version sounds best to you.

In Part II, you will hear examples of a drum and cymbal sound hitting simultaneously or near-simultaneously. In each example, you will be asked to determine whether the drum hits first, the cymbal hits first or whether they hit at the same time.

In Part III, you will hear grouping of two bass drum hits. Each group has two pairs of hits, spaced differently in time. You will be asked to determine for each group which pair is spaced farther apart.

You will be given oral instructions before each Part. If it is not clear enough at that time what you are being asked to do, you may ask questions then.

Thank you very much for your participation!

Part I

Instructions:

You are about to hear pairs of short drum rhythms. Each pair contains alternate versions of the same rhythm. You are asked to answer which version sounds best to you, or whether the versions sound the same. When doing this, use your feelings and sense of rhythm. It is not necessary to think it over, there is no right or wrong answer. For each musical example, respond by placing an "x" after either "First Version," "Second Version," or "No Difference," according to your preference.

1. Which sounded best?

First Version ____ Second Version ____ No Difference ____

2. Which sounded best?

First Version ____ Second Version ____ No Difference ____

3. Which sounded best?

First Version ____ Second Version ____ No Difference ____

4. Which sounded best?

First Version ____ Second Version ____ No Difference ____

5. Which sounded best?

First Version ____ Second Version ____ No Difference ____

6. Which sounded best?

First Version ____ Second Version ____ No Difference ____

7. Which sounded best?

First Version _____ Second Version _____ No Difference _____

8. Which sounded best?

First Version _____ Second Version _____ No Difference _____

Part II

Instructions:

You are about to hear examples of a drum and cymbal sound hitting simultaneously or near-simultaneously. Each example will repeat six times, to help you in deciding what the order is. For each example, answer whether you think the drum hit first, the cymbal hit first, or whether they hit at the same time. For each musical example, respond by placing an "x" under either "Drum Hit First," "Cymbal Hit First," or "Simultaneous Hit," according to your preference.

1. Drum Hit First Cymbal Hit First Simultaneous Hit

2. Drum Hit First Cymbal Hit First Simultaneous Hit

3. Drum Hit First Cymbal Hit First Simultaneous Hit

Part III

Instructions:

You are about to hear groupings of two bass drum hits. Each group has two pairs of hits, spaced differently in time. Answer for each group which pair is spaced farther apart in time. For each group, the pairs will be repeated six times to help you in deciding which pair is spaced farther apart. Answer by placing an "x" beneath your choice.

1. Which pair is spaced farther apart in time?

First Pair

Second Pair

2. Which pair is spaced farther apart in time?

First Pair

Second Pair

3. Which pair is spaced farther apart in time?

First Pair

Second Pair

4. Which pair is spaced farther apart in time?

First Pair

Second Pair

Background Information

1. Your age is ____ years.
2. Are you male or female? (place an "x" after your sex)
 Female ____ Male ____
3. Have you ever had any musical training, either in school or in private lessons?
 yes ____ no ____
4. If so, which instrument(s) did you study?

5. If so, how many years have you studied some musical instrument?
 ____ years.
6. Do you currently play any musical instruments?
 yes ____ no ____
7. If you currently play any musical instruments, write on the line the instrument(s) you play.

8. If you currently play musical instruments, write on the line how many years of experience you have with those instruments.
 ____ years.
9. For music you like, are you in favor of musicians using drum machines? (circle the number which best represents your preference)
 In Favor 1 2 3 4 5 6 7 Opposed To
10. How important is it to you to listen to music on a high quality stereo system? (circle the number which best represents your preference)
 Very Important 1 2 3 4 5 6 7 Not Important
11. About how many hours a day do you spend listening to music? ____ hours.

This ends the study. Thanks again for participating.

LIST OF REFERENCES

- Bengtsson, Ingmar and Gabrielsson, Alf. "Rhythm Research in Uppsala." Music, Room and Acoustics: Publications Issued by The Royal Swedish Academy of Music 17 (1977): 19-56.
- Bengtsson, Ingmar and Gabrielsson, Alf. "Methods For Analyzing Performance of Musical Rhythm." Scandinavian Journal of Psychology 21 (1980): 257-68.
- Boom, Michael. Music Through MIDI: Using MIDI to Create Your Own Electronic Music System. Redmond: Microsoft Press, 1987.
- Boulanger, Richard. "Conducting the MIDI Orchestra, Part 1: Interviews with Max Mathews, Barry Vercoe, and Roger Dannenberg." Computer Music Journal Summer 1990: 34-39.
- De Furia, Steve. The MIDI Book: Using MIDI and Related Interfaces. With Joe Scacciaferro. Rutherford: Third Earth Publications, Inc.: 1986.
- Friberg, Anders., et al. "Performance Rules For Computer-Controlled Contemporary Keyboard Music." Computer Music Journal Summer 1991: 49-55.
- Gabrielsson, Alf. "Performance of Rhythm Patterns." Scandinavian Journal of Psychology 15 (1974): 63-72.
- Gabrielsson, Alf. "Experimental Research on Rhythm." Humanities Association Review 1979: 39-62.
- Haydon, Glenn. Introduction to Musicology: A Survey of the Fields, Systematic and Historical, of Musical Knowledge and Research. New York: Prentis-Hall, Inc., 1941.
- Michon, J. A. "Differential Sensitivity in the Perception of Repeated Temporal Intervals." Acta Psychologica 22 (1964): 441-50.
- Milano, Dominic, ed. Mind Over MIDI. Milwaukee: H. Leonard Books, 1987.
- Rubenking, Janet. "Add A Musical Dimension to Your PC with MIDI." PC Magazine 12 March 1991: 355-366.
- Sears, C. H. "A Contribution to the Psychology of Rhythm." American Journal of Psychology 13 (1902): 28-61.
- Seashore, C. E. Psychology of Music. New York: McGraw-Hill, 1938.

Smith, Leland. "SCORE--A Musician's Approach to Computer Music." Journal of the Acoustical Society of America 20 (1972): 7-14.

Woodrow, H. "A Quatitative Study of Rhythm." Archives of Psychology 14 (1909).

General References

- Anderson, David P. "Accurately Timed Generation of Discreet Musical Events." Computer Music Journal Fall 1986: 48-55.
- Anderton, Craig. MIDI For Musicians. New York: Amsco Publications, 1986.
- Friberg, Anders. "Generative Rules For Music Performance: A Formal Description of a Rule System." Computer Music Journal Summer 1991: 56-71.
- Gabrielsson, Alf. "Similarity Ratings and Dimension Analyses of Auditory Rhythm Patterns II." Scandinavian Journal of Psychology 14 (1973): 161-76.
- Garner, W. R. and Miller, G. A. "Differential Sensitivity to Intensity as a Function of the Duration of the Comparison Tone." Journal of Experimental Psychology 34 (1944): 450-463.
- Garner, W. R. and Miller, G. A. "The Masked Threshold of Pure Tones as a Function of Duration." Journal of Experimental Psychology 37 (1947): 293-303.
- Henry, Franklin M. "Discrimination of the Duration of a Sound." Journal of Experimental Psychology 38 (1948): 735-43.
- Jaffe, David. "Ensemble Timing in Computer Music." Computer Music Journal Winter 1985: 38-48.
- Loy, Gareth. "Musicians Make a Standard: The MIDI Phenomenon." Computer Music Journal Winter 1985: 8-26.
- Moore, Richard F. "The Dysfunctions of MIDI." Computer Music Journal Spring 1988: 19-28.
- Povel, Dirk-Van. "Temporal Structure of Performed Music: Some Preliminary Observations." Acta Psychologica 41 (1977): 309-20.
- Rona, Jeff. MIDI, The Ins, Outs and Thrus. Ed. Ronny S. Schiff. Milwaukee: H. Leonard Books, 1987.
- Scholes, Percy A. The Oxford Companion To Music, 10th ed. Ed. John Owen Ward. New York: Oxford University Press, 1970.
- Small, A. M. and Campbell, R. A. "Temporal Differential Sensitivity for Auditory Stimuli." American Journal of Psychology 75 (1962): 401-10.

Tove, P.A., et al. "Direct-recording Frequency and Amplitude Meter For Analysis of Musical and other Sonic Wave Forms." Journal of the Acoustical Society of America 39 (1966): 362-71.

Zicarelli, David. "M and Jam Factory." Computer Music Journal Winter 1987: 13-29.

MICHIGAN STATE UNIV. LIBRARIES



31293008994224